NOTE TO USERS

This reproduction is the best copy available.

UMI

**University of Alberta**

# A putative copper-transporting P-type ATPase from *Brassica napus* complements the *Saccharomyces cerevisiae ccc2* mutant

by

Jennafer Lynn Southron

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Plant Biology

Department of Biological Sciences

Edmonton, Alberta

Spring 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-69764-9

Canada

**University of Alberta**

**Library Release Form**

**Name of Author:** Jennafer Lynn Southron

**Title of Thesis:** A putative copper-transporting P-type ATPase from *Brassica napus* complements the *Saccharomyces cerevisiae ccc2* mutant

**Degree:** Master of Science

**Year this Degree Granted:** 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Jennafer Southron

52113A Range Road 275
Spruce Grove, Alberta
T7X 3V2 Canada

Date: January 30, 2002

**University of Alberta**

**Faculty of Graduate Studies and Research**

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **A putative copper-transporting P-type ATPase from *Brassica napus* complements the *Saccharomyces cerevisiae ccc2* mutant** submitted by Jennafer Lynn Southron in partial fulfillment of the requirements for the degree of Master of Science in Plant Biology.

Gregory J. Taylor (supervisor)

Allen G. Good

William J. Page

Janusz Zwiazek

Date: January 29, 2002

# Abstract

The Type IB subfamily of the P-type ATPases ($P_{1B}$-ATPases) contains at least two groups, the $Cu^{2+}$- and $Cd^{2+}$-ATPases, that differ in substrate specificity. When this study was initiated, there was no information about the functional role of $P_{1B}$-ATPases in plants. A search of the literature and sequence databases revealed four *Arabidopsis* sequences that encode putative $P_{1B}$-ATPases. In this study, a *Brassica napus* cDNA encoding a homologue of one of these database sequences was cloned. Based on sequence analysis, the putative $P_{1B}$-ATPase encoded by this cDNA was predicted to transport copper. A *Saccharomyces cerevisiae* complementation assay demonstrated that the *B. napus* cDNA mediated complete rescue of a mutant deficient in a copper-transporting $P_{1B}$-ATPase, providing functional evidence in support of a role in copper transport. The current work is amongst the first to demonstrate that a $P_{1B}$-ATPase is functional in plants and provides information about its putative substrate specificity.

# Acknowledgments

Many thanks are extended to my supervisor, Dr. Gregory Taylor, for his support and guidance. I would also like to extend my gratitude to the members of my supervisory committee, Dr. Allen Good and Dr. William Page, as well as my external examiner, Dr. Janusz Zwiazek, for their time, effort, and discussion. A special thank you is extended to the laboratories of Dr. Isobel Parkin, for the gift of the *Brassica napus* cDNA library, and Dr. Diane Cox, for the *Saccharmoyces cerevisiae* strains and pG3 vectors. I am grateful for the assistance and friendship provided by all members, past and present, of the Taylor lab. In particular I would like to acknowledge the guidance provided by Dr. Urmila Basu, Julie Stephens, and Valar Anoop. I gratefully acknowledge the support provided to me by the Department of Biological Sciences, in the form of a Teaching Assistantship (1998-1999; 2000-2001), and the University of Alberta/Province of Alberta, in the form of a Province of Alberta Graduate Scholarship (1998-1999). I am also grateful for the support that was provided by the National Science and Engineering Research Council in the form of a PGS-A research scholarship (1999-2000).

# Table of Contents


**1 General introduction** ..... 1

1.1 The P-type ATPase superfamily of transport proteins ..... 1

1.2 Structure and reaction cycle of the P-type ATPases ..... 2

1.3 The P-type ATPases cluster according to their substrate specificity ..... 3

1.4 The $P_{1B}$-ATPases ..... 6

1.5 The *Saccharomyces cerevisiae CCC2* gene encodes a $Cu^{2+}$-ATPase that plays an important role in copper and iron homeostasis ..... 9

1.6 Specific components of the copper homeostasis pathway are required for the delivery of copper to the Ccc2 and Fet3 proteins ..... 9

1.7 The Fet3, Ftr1, Fre1, and Fre2 proteins are required for high-affinity iron uptake ..... 11

1.8 Regulation of high-affinity iron and copper uptake ..... 11

1.9 Low-affinity transport systems function under metal-replete conditions ..... 13

1.10 Metal supplementation can rescue growth defects that are due to a defective high-affinity iron uptake system ..... 13

1.11 *Saccharomyces cerevisiae* strains other than the *ccc2* mutant may be of use for studying $P_{1B}$-ATPases ..... 15

1.12 $P_{1B}$-ATPases in plants? ..... 16

1.13 Literature cited ..... 21

**2 The identification, analysis, and cloning of plant sequences corresponding to putative $P_{1B}$-ATPases** ..... 28

2.1 Introduction ..... 28

2.2 Materials and methods ..... 30

2.2.1 The identification and analysis of plant sequences that correspond to putative $P_{1B}$-ATPases ..... 30

2.2.1.1 Literature search....30

2.2.1.2 Sequence database search: Basic Local Alignment Search Tool at the National Center for Biotechnology Information....30

2.2.1.3 Analysis of sequences that were identified in the literature and sequence database searches....32

2.2.1.3.1 Identification of P-type and Type IB motifs....32

2.2.1.3.2 Predicting substrate specificity....33

2.2.1.3.3 Grouping expressed sequence tags and full-length sequences....33

2.2.1.3.4 Hydrophobicity....33

2.2.2 Cloning of a *Brassica napus* cDNA that encodes a $P_{1B}$-ATPase....34

2.2.2.1 General techniques....34

2.2.2.1.1 Isolation of nucleic acids from plant tissue....34

2.2.2.1.2 Polymerase chain reaction primers....35

2.2.2.1.3 Polymerase chain reactions....35

2.2.2.1.4 Agarose gels....37

2.2.2.1.5 Gel extraction....37

2.2.2.1.6 Production of a T-vector for cloning polymerase chain reaction products....38

2.2.2.1.7 Ligations....38

2.2.2.1.8 The production of competent DH5∝ *Escherichia coli* cells....38

2.2.2.1.9 Transformation of competent DH5∝ *Escherichia coli* cells....39

2.2.2.1.10 Preparation of plasmid DNA....40

2.2.2.1.11 Restriction enzyme digest reactions....40

2.2.2.1.12 Sequencing....41

2.2.2.1.13 The identification of plasmids harboring an appropriate insert....42

2.2.2.1.14 Glycerol stocks....42

2.2.2.1.15 $^{32}$P-labeling of DNA fragments....42

2.2.2.2 Amplification of polymerase chain reaction products from plant putative $P_{1B}$-ATPases....43

2.2.2.3 Production of polymerase chain reaction-based probes that will be used to screen a *Brassica napus* cDNA library....44

2.2.2.4 Screening the amplified *Brassica napus* cDNA library....45

2.2.2.5 Isolation of the 5'-coding sequence that is missing from the ~2000 bp library clone by reverse transcriptase-polymerase chain reaction....46

2.2.2.6 Production of a full-length *Brassica napus* cDNA....47

2.2.2.7 Sequencing and analysis of the full-length *Brassica napus* cDNA....47

2.3 Results....48

2.3.1 Identification and analysis of plant sequences that correspond to putative $P_{1B}$-ATPases....48

2.3.1.1 Literature search....48

2.3.1.2 Sequence database search: Basic Local Alignment Search Tool at the National Center for Biotechnology Information....49

2.3.1.3 Analysis of sequences that were identified in the literature and sequence database searches....52

2.3.1.3.1 Identification of P-type and Type IB motifs and the prediction of substrate specificity....52

2.3.1.3.2 Grouping expressed sequence tags and full-length sequences ..... 54

2.3.1.3.3 Hydrophobicity ..... 56

2.3.2 Cloning of a *Brassica napus* cDNA that encodes a $P_{1B}$-ATPase ..... 58

2.3.2.1 The amplification of polymerase chain reaction products from plant putative $P_{1B}$-ATPases ..... 58

2.3.2.2 Production of polymerase chain reaction-based probes and the screening of an amplified *Brassica napus* cDNA library ..... 64

2.3.2.3 Isolation of the 5'-coding sequence that is missing from the ~2000 bp library clone by reverse transcriptase-polymerase chain reaction and production of a full-length cDNA ..... 68

2.3.2.4 Sequencing and analysis of the full-length *Brassica napus* cDNA ..... 70

2.4 Discussion ..... 73

2.5 Literature cited ..... 107

**3 Functional complementation of the *Saccharomyces cerevisiae ccc2* mutant by a *Brassica napus* cDNA (AY045772/15636780) encoding a putative copper-transporting P-type ATPase** ..... 111

3.1 Introduction ..... 111

3.2 Materials and methods ..... 113

3.2.1 Construction of the yeast expression vector ..... 113

3.2.2 General growth conditions and techniques ..... 115

3.2.3 *Saccharomyces cerevisiae* strains ..... 115

3.2.4 *Saccharomyces cerevisiae* transformation ..... 116

3.2.5 Iron-limited, iron-sufficient, copper-sufficient, and induction media ..... 116

3.2.6 Growth curves: carbon source ..... 117

3.2.7 Isolation of total RNA and Northern hybridization ..... 118

3.2.8 Plating assay ..... 118

3.2.9 Growth curves: complementation assay ..... 119

3.3 Results ..... 120

3.3.1 Growth curves: carbon source ..... 120

3.3.2 Plating assay ..... 121

3.3.3 Growth Curves: complementation assay ..... 125

3.4 Discussion ..... 126

3.5 Literature cited ..... 136

**4 Concluding discussion** ..... 138

4.1 Literature cited ..... 145

**5 Appendix** ..... 148

# List of Tables

2.1 Polymerase chain reaction (PCR) primer sequences ..... 104

2.2 Polymerase chain reaction products that could be produced by primers based on the DKTGT[LIVM][TIS] and GDGxNDxP amino acid sequences..... 105

2.3 Multiple alignment results..... 106

# List of Figures


1.1 The P-type ATPase superfamily ..... 19

1.2 Model illustrating the role Ccc2p plays in copper and iron homeostasis of *S. cerevisiae* ..... 20

2.1 The position of the PCR-S and PCR-A primers relative to the *RAN1* (A) and *PAA1* (B) coding sequences ..... 79

2.2 The position of the primers and probes that were produced in preparation to screen the amplified *B. napus* library for cDNAs that encode RAN1 (A) or PAA1 (B) homologues ..... 80

2.3 The position of the primers and a probe that were produced in preparation to screen the amplified *B. napus* library for cDNA(s) that encode a HMA1 homologue ..... 81

2.4 The position of the primers used in the RT-PCR strategy that was intended to amplify the 5'-coding region that corresponds to the ~2000 bp cDNA ..... 82

2.5 Plasmids that were used in the ligation reaction to produce a full-length *B. napus* cDNA ..... 83

2.6 The full-length *B. napus* cDNA in the pBluescript SK- plasmid (pBluescript-cDNA) ..... 84

2.7 The position of primers that were used to sequence the *B. napus* cDNA relative to the *B. napus* coding sequence ..... 85

2.8 Phylogenetic tree ..... 86

2.9 Expressed sequence tags (ESTs) that correspond to the *RAN1* (A) and *HMA1* (B) coding sequences ..... 87

2.10 Expressed sequence tags (ESTs) that correspond to the *HMA4* coding sequence ..... 88

2.11 Hydrophobicity profiles for PAA1 (A), RAN1 (B), HMA1 (C), and HMA4 (D) ..... 89

2.12 Hydrophobicity profiles for Ccc2p (A; $P_1$-ATPase), Pma2p (B; $P_2$-ATPase), and KdpB (C; $P_3$-ATPase) ..... 90

2.13 Polymerase chain reaction (PCR) products amplified using the PCR-S/PCR-A primer pair ..... 91

2.14 The partial coding sequence (CDS) of PCR-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence ..... 92

2.15 The partial coding sequence (CDS) of PCR-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence ..... 93

2.16 The partial sequence of PCR-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* genomic sequence ..... 94

2.17 The partial sequence of PCR-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* genomic sequence ..... 95

2.18 The position of the probes that were produced to screen the amplified *B. napus* library for cDNA(s) that encode a RAN1 homologue and the partial clones that were isolated using Probe1 ..... 96

2.19 The Probe3-Bn1 and Probe3-Bn2 sequences aligned with the appropriate region of the *RAN1* genomic or coding sequence ..... 97

2.20 Reverse transcriptase-polymerase chain reaction (RT-PCR) products ..... 98

2.21 The partial sequence of RT-Bn1 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer ..... 99

2.22 The partial sequence of RT-Bn2 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer ..... 100

2.23 The partial sequence of RT-Bn3 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer ..... 101

2.24 Hydrophobicity profile for the protein encoded by the *B. napus* cDNA ..... 102

2.25 The protein encoded by the *B. napus* cDNA aligned with the updated RAN1 amino acid sequence (AAC79141/6850337) ..... 103

3.1 The full-length *B. napus* cDNA in the pBluescript SK- plasmid (pBluescript-cDNA) ..... 129

3.2 The full-length *B. napus* cDNA in the pYES3 vector (pYES3-cDNA).....130

3.3 Carbon source growth curves.....131

3.4 Complementation assay plates after a short (72 hour) incubation.....132

3.5 Transcript levels from the *B. napus* cDNA.....133

3.6 Complementation assay plates after an extended (240 hour) incubation.....134

3.7 Complementation assay growth curves.....135

5.1 The partial sequence of PCR-At1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* genomic sequence.....148

5.2 The partial sequence of PCR-At1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* genomic sequence.....149

5.3 The partial coding sequence (CDS) of PCR-At2 that was produced with the T7 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.....150

5.4 The partial coding sequence (CDS) of PCR-At2 that was produced with the T3 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.....151

5.5 The partial coding sequence (CDS) of PCR-Bn2 that was produced with the T3 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.....152

5.6 The partial coding sequence (CDS) of PCR-Bn2 that was produced with the T7 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.....153

5.7 The partial sequence of the ~2000 bp cDNA that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).....154

5.8 The partial sequence of the ~2000 bp cDNA that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).....155

5.9 The partial sequence of the ~1000 bp cDNA that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).....156

5.10 The Probe3-Bn1 sequence aligned with the appropriate region of the *RAN1* genomic or coding sequence....157

5.11 The Probe3-Bn2 sequence aligned with the appropriate region of the *RAN1* genomic or coding sequence....158

5.12 The Probe4-Bn1 sequence aligned with the appropriate region of the *PAA1* genomic or mRNA/CDS....159

5.13 The Probe5-At1 sequence aligned with the appropriate region of the *HMA1* genomic or coding sequence....160

5.14 The Probe6-At1 sequence aligned with the appropriate region of the *HMA1* genomic sequence....161

5.15 The Probe6-Bn1 sequence aligned with the appropriate region of the *HMA1* genomic sequence....162

5.16 The Probe6-Bn1 coding sequence (CDS) aligned with the appropriate region of the *HMA1* coding sequence....163

5.17 The partial coding sequence of RT-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS)....164

5.18 The partial coding sequence of RT-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS)....165

5.19 The partial coding sequence of RT-Bn2 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS)....166

5.20 The partial coding sequence of RT-Bn2 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS)....167

# List of Abbreviations

| | |
|---|---|
| AGI | *Arabidopsis* Genome Initiative |
| ATP | adenosine triphosphate |
| ATPase | adenosine triphosphatase |
| BLAST | Basic Local Alignment Search Tool |
| bp | base pair |
| cDNA | complementary DNA |
| CDS | coding sequence |
| C-terminal; C-terminus | carboxyl-terminal; carboxyl-terminus |
| DDBJ | DNA Database of Japan |
| °C | degrees Celsius |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| dNTP | deoxynucleotide triphosphate |
| dTTP | deoxythymidine triphosphate |
| EMBL | European Molecular Biology Laboratory |
| EST | expressed sequence tag |
| x g | multiplied by the standard acceleration of gravity |
| GI | GenInfo |
| IPTG | isopropylthio-β-D-galactoside |
| LB | Luria-Bertani |
| LBA | Luria-Bertani ampicillin |
| MES | 2-[*N*-morpholino]ethanesulfonic acid |

| | |
|---|---|
| µl; ml | microliter; milliliter |
| µM; mM; M | micromolar; millimolar; molar |
| mm | millimeter |
| MIPS | Munich Information Center for Protein Sequences |
| mRNA | messenger RNA |
| ng; µg | nanogram; microgram |
| NCBI | National Center for Biotechnology Information |
| N-terminal; N-terminus | amino-terminal/amino-terminus |
| $OD_{600}$ | optical density at 600 nanometers |
| $OPA^{+}$ | One-Phor-All Buffer PLUS |
| PCR | polymerase chain reaction |
| PDB | Protein Data Bank |
| PHI-BLAST | Pattern-Hit Initiated Basic Local Alignment Search Tool |
| PIR | International Protein Sequence Database |
| pmoles | picomoles |
| PRF | Protein Research Foundation |
| RNA | ribonucleic acid |
| RPM | revolutions per minute |
| RT-PCR | reverse transcriptase-polymerase chain reaction |
| SDS-PAGE | sodium dodecyl sulphate-polyacrylamide gel electrophoresis |
| SERCA | sarco(endo)plasmic reticulum $Ca^{2+}$-ATPase |
| TBE | tris-borate/EDTA |
| TCO | *trans*-cyclooctene |

| | |
|---|---|
| TE | tris-EDTA |
| TIGR | The Institute for Genomic Research |
| X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactoside |
| YNB | yeast nitrogen base |

## Amino acids

| | |
|---|---|
| A | alanine |
| C | cysteine |
| D | aspartate |
| E | glutamate |
| F | phenylalanine |
| G | glycine |
| H | histidine |
| I | isoleucine |
| K | lysine |
| L | leucine |
| M | methionine |
| N | asparagine |
| P | proline |
| Q | glutamine |
| R | arginine |
| S | serine |
| T | threonine |
| V | valine |

| | |
|---|---|
| W | tryptophan |
| Y | tyrosine |
| x | any amino acid |

## Nucleic acids

| | |
|---|---|
| A | adenine |
| C | cytosine |
| G | guanine |
| T | thymine |

## Elements

| | |
|---|---|
| Ag | silver |
| Ca | calcium |
| Cd | cadmium |
| Cl | chlorine |
| Co | cobalt |
| Cu | copper |
| Fe | iron |
| H | hydrogen |
| K | potassium |
| Mg | magnesium |
| Mn | manganese |
| Na | sodium |
| Ni | nickel |
| P | phosphorus |

| | |
|---|---|
| Pb | lead |
| Zn | zinc |

# 1 General introduction

## 1.1 The P-type ATPase superfamily of transport proteins

The P-type ATPase superfamily is a group of proteins that are involved in the transport of a variety of charged substrates across biological membranes. This transport is ATP-dependent and involves the phosphorylation of an aspartate (D) residue found in the amino acid sequence DKTGT[LIVM][TIS] (Prosite PS00154). The presence of these amino acids is the only requirement for a polypeptide to be classified as a putative P-type ATPase. Nonetheless, additional amino acid sequence motifs are conserved in most P-type ATPases and mutational analysis has been used to propose a functional role for some of these motifs (Fagan and Saier, 1994; Moller *et al.*, 1996; Solioz and Vulpe, 1996; Axelsen and Palmgren, 1998). The P-type ATPase reaction cycle (Section 1.2) has been proposed to involve at least two conformational states. The conserved amino acid sequence TGE is believed to participate in transducing the energy of ATP hydrolysis to conformational change(s) in the protein (energy transduction; Moller *et al.*, 1996). The GDGxNDxP amino acid motif (where x represents any amino acid) has been proposed to function in ATP-binding (Fagan and Saier, 1994) or as a hinge that affords the protein the flexibility required to undergo conformational change (Moller *et al.*, 1996). An invariant proline (P) is found approximately 43 amino acids upstream of the DKTGT[LIVM][TIS] motif and is thought to function in ion transduction (Solioz and Vulpe, 1996). The order of these motifs, which presumably specifies their spatial relationship, has been maintained in all P-type ATPases (Fagan and Saier, 1994). In addition to these motifs, eight dispersed regions that exhibit a high level of sequence similarity have been identified (Axelsen and Palmgren, 1998). These regions were identified as sequences that always aligned and contained only minimal deletions in partial multiple alignments of 211 P-type ATPases. These motifs and conserved regions represent a small fraction of the amino acid sequence of P-type ATPases; the overall sequence similarity of the P-type ATPases is actually low. The motifs and conserved regions are, however, sufficient to suggest that the P-type ATPases share rudimentary structural and mechanistic features.

## 1.2 Structure and reaction cycle of the P-type ATPases

The amino acids that are conserved in most P-type ATPases have been proposed to participate in functions or structures that play a central role in the catalytic mechanism of transport (Silver *et al.*, 1989; Fagan and Saier, 1994; Lutsenko and Kaplan, 1995; Moller *et al.*, 1996; Axelsen and Palmgren, 1998). General information that is gained from an individual P-type ATPase and involves the conserved motifs or regions of high sequence similarity can, therefore, be extended to other P-type ATPases. The P-type ATPases have also been proposed to share a common core structure (Section 1.3; Lutsenko and Kaplan, 1995), thus structural information from individual proteins can also be extended to other P-type ATPases. While many P-type ATPases function as a single subunit, some P-type ATPases require additional subunits (Section 1.3). When a P-type ATPase is composed of multiple subunits, the common P-type ATPase features are usually found within the main catalytic subunit.

General features of the catalytic cycle developed for the sarco(endo)plasmic reticulum $Ca^{2+}$-ATPase (SERCA; MacLennan *et al.*, 1997; MacLennan and Green, 2000; Toyoshima *et al.*, 2000; Lee and East, 2001) provide information about how proteins from other branches of the P-type ATPase superfamily (Silver *et al.*, 1989; Solioz and Vulpe, 1996; Voskoboinik *et al.*, 2001) may function. The first conformational state, referred to as the E1 state, exhibits a high affinity for ATP and the substrate that is to be transported. The ATP is likely to bind to a common ATP/nucleotide-binding region located on the catalytic subunit. The substrate-binding sites will likely vary with the type of substrate and in the case of $P_{1A}$-ATPases may be on a separate subunit (Section 1.3; Palmgren and Axelsen, 1998). When the substrate and ATP are bound, the P-type ATPase can be phosphorylated by the γ-phosphate of the bound ATP. This phosphorylation occurs at the invariant aspartate (D) in the phosphorylation site (DKTGT[LIVM][TIS]). After phosphorylation, the P-type ATPase undergoes a rate-limiting transition to the second conformational state (E2) that is accompanied by the loss of the transported substrate. Hydrolysis of the bound phosphate regenerates the substrate-

binding site(s) and returns the P-type ATPase to the E1 state allowing the reaction cycle to begin anew.

A number of structural and biochemical studies have examined the structure of the SERCA proteins (Toyoshima *et al.*, 1993; MacLennan *et al.*, 1997; Zhang *et al.*, 1998; MacLennan *et al.*, 2000; Toyoshima *et al.*, 2000). The majority of the protein, including the N- and C-termini, is located in the cytoplasm, a small section is in the membrane, and an even smaller portion of the protein is in the lumen. The overall shape of this $Ca^{2+}$-ATPase can be described as a large cytoplasmic headpiece connected to the transmembrane domain and a limited lumenal domain by a short stalk segment. The cytoplasmic headpiece consists of phosphorylation, nucleotide-binding, and actuator domains. Flexible connections are found between the phosphorylation and nucleotide-binding domains, and between the actuator domain and the rest of the protein. This flexibility is thought to play an important role in the conformational changes that occur during the reaction cycle. It has also been proposed that the actuator domain is involved in the transmission of the major conformational changes. The transmembrane region contains six transmembrane alpha-helices that are thought to be found in all P-type ATPases and four transmembrane alpha-helices that are only found in a subset of the P-type ATPases. These two groups of transmembrane alpha-helices are segregated from one another and are, therefore, consistent with the core structure of six transmembrane domains that was proposed by Lutsenko and Kaplan (1995; Section 1.3). The structure of SERCA proteins is consistent with the putative mechanism outlined above (Toyoshima *et al.*, 2000; Lee and East, 2001).

## 1.3 The P-type ATPases cluster according to their substrate specificity

Although a number of similarities can be identified in all P-type ATPases, the overall similarity is low and functional and mechanistic differences are apparent. Detailed amino acid analysis and alignments can be used to divide the P-type ATPase

superfamily into smaller divisions and to identify features that are characteristic of these divisions. A number of phylogenetic studies have revealed that the P-type ATPases tend to cluster according to the substrate(s) they transport (Fagan and Saier, 1994; Lutsenko and Kaplan, 1995, Moller *et al.*, 1996; Axelsen and Palmgren, 1998).

The work of Axelsen and Palmgren (1998) represents the most recent and extensive phylogenetic analysis of the P-type ATPase superfamily and the division outlined by these authors will be used throughout this study. The low overall similarity of the P-type ATPases made a complete multiple alignment of numerous full-length P-type ATPase amino acid sequences impossible (Axelsen and Palmgren, 1998). Partial alignments, however, were possible and revealed a core sequence with minimal deletions that exhibits a high level of sequence similarity (Axelsen and Palmgren, 1998). This core sequence was identified in 159 full-length, P-type ATPase sequences and used in a phylogenetic analysis. The P-type ATPase superfamily was divided into five major families (Type 1-5; Axelsen and Palmgren, 1998; Figure 1.1). The Type 1, 2, and 3 families were further divided into the Type 1A and 1B; Type 2A, 2B, 2C, and 2D; and Type 3A and 3B subfamilies, respectively (Figure 1.1). Some of the subfamilies were subsequently divided into different groups (Figure 1.1). The families, subfamilies, and groups identified by Axelsen and Palmgren (1998) clustered according to substrate specificity. The Type 1A subfamily contains KbpB proteins that are involved in $K^+$-transport. At least two groups, the $Cu^{2+}$- and $Cd^{2+}$-ATPases, are found within the Type 1B subfamily. The Type 2A and 2B subfamily members are involved in the transport of $Ca^{2+}$. The $Na^+/K^+$- and $H^+/K^+$-ATPase groups make up the Type 2C subfamily. The Type 2D subfamily and Type 5 family contain ATPases of unknown function. The plasma membrane $H^+$-ATPases and the $Mg^{2+}$-ATPases compose the Type 3A and Type 3B subfamilies, respectively. Members of the Type 4 family transport aminophospholipids. It should be noted that these divisions can be designated in a variety of ways. As an example, the terms Type 1B, P-type ATPase; Type 1B ATPase; and $P_{1B}$-ATPase all describe a member of the Type 1B subfamily.

The various P-type ATPase families, subfamilies, and groups (Axelsen and Palmgren, 1998) not only coincide with substrate specificity, but also with a number of other attributes. For example, amino acid motifs that are unique to a single division of the P-type ATPase superfamily can be identified (Section 1.4; Solioz and Vulpe, 1996). Furthermore, although P-type ATPases are ubiquitous, particular families or subfamilies may exhibit a limited distribution (Axelsen and Palmgren, 1998; Palmgren and Axelsen, 1998). The $P_{1A}$- and $P_{3B}$-ATPases have only been described in bacteria, while the $P_{2B}$-, $P_{2C}$-, $P_{2D}$-, $P_4$-, and $P_5$-ATPases have only been identified in eukaryotes. The $P_3$-ATPases have been found in archaebacteria, plants, and fungi. The $P_{1B}$- and $P_{2A}$-ATPases have been discovered in bacteria, archaebacteria, and eukaryotes. As new P-type ATPases are identified, the distribution of a particular division may be broadened.

Some families or subfamilies of the P-type ATPase superfamily require multiple subunits to mediate transport (Palmgren and Axelsen, 1998). The $P_{1A}$-ATPases (KdpB-ATPases) require a total of three subunits: the catalytic subunit (KdpB) that hydrolyzes ATP, a subunit (KdpA) that binds the substrate, and a subunit (KdpC) that may be involved in stabilizing the complex between the first two subunits (KdpB and KdpA). It has been proposed that the $P_{1A}$-ATPases represent an ancestral P-type ATPase (Axelsen and Palmgren, 1998). In more recent P-type ATPases, the ATP hydrolyzing and substrate binding subunits have become fused (Axelsen and Palmgren, 1998). The $P_{2C}$-ATPases ($Na^+/K^+$- and $H^+/K^+$-ATPases) require a main catalytic subunit ($\alpha$) and an additional subunit ($\beta$), which has been proposed to represent a remnant of the KdpC subunit (Axelsen and Palmgren, 1998). The $Na^+/K^+$-ATPases also associate with a third subunit ($\gamma$) that seems to play an important role in $K^+$ activation of these ATPases (Palmgren and Axelsen, 1998). The $P_{2A}$-ATPases interact with regulatory proteins and the $P_{3A}$-ATPases associate with proteins of unknown function (Palmgren and Axelsen, 1998). In other divisions, such as the Type IIB subfamily, regulatory regions are incorporated into the N- or C-terminus of the main catalytic subunit and may represent a fusion of subunits (Palmgren and Axelsen, 1998).

Prior to the work of Axelsen and Palmgren (1998), other researchers had divided the P-type ATPase superfamily into groups that correlated (although to a lesser extent) with substrate specificity (Fagan and Saier, 1994; Lutsenko and Kaplan, 1995; Moller *et al.*, 1996). The divisions established in most of these studies were based on the alignment of a limited number of amino acid sequences and will not be discussed further. The work of Lutsenko and Kaplan (1995) is important because of the unique method used to group amino acid sequences. Lutsenko and Kaplan (1995) examined the hydropathy profiles for a number of P-type ATPases. A core structure consisting of six transmembrane domains was identified in all P-type ATPases and unique additions to this core sequence were identified and used to group the P-type ATPases into three divisions ($P_1$-, $P_2$-, and $P_3$-ATPases). The divisions established by Lutsenko and Kaplan (1995) correlated with substrate specificity. The $P_1$-ATPases were reported to be heavy metal transporters, the $P_2$-ATPases transported non-heavy metals, and the $P_3$-ATPases represented the $K^+$-transporting, KdpB-ATPases. The core structure was proposed to function in aspects of the transport mechanism that were shared by all P-type ATPases, while the unique regions were thought to satisfy specific requirements of the different substrates. The core structure consists of a pair of transmembrane domains before the TGE motif, a transmembrane pair between the TGE motif and the phosphorylation site (DKTGT[LIVM][TIS]), and another set of two transmembrane domains after the GDGxNDxP sequence. A hydrophobic region was identified before the TGE motif of the $P_1$-ATPases and was proposed to correspond to two transmembrane domains that are not found in the core structure. An additional hydrophobic region was identified C-terminal of the core sequence of the $P_2$- and $P_3$-ATPases. The C-terminal, hydrophobic region found in the $P_2$-ATPases is longer than that of the $P_3$-ATPases (Lutsenko and Kaplan, 1995).

## 1.4 The $P_{1B}$-ATPases

The $P_1$-ATPase division established by Lutsenko and Kaplan (1995) is equivalent to the $P_{1B}$-ATPase subfamily identified by Axelsen and Palmgren (1998). As a result, the hydropathy profile outlined for the $P_1$-ATPases (Section 1.3; Lutsenko and Kaplan,

1995), which is distinct from the hydropathy profiles of the other members of the P-type ATPase superfamily, can be applied to the $P_{1B}$-ATPases. Regions of conserved amino acids have also been identified that are found in the $P_{1B}$-ATPases, but are not found in other members of the P-type ATPase superfamily (Nucifora *et al.*, 1989; Silver *et al.*, 1989; Tanzi *et al.*, 1993; Solioz and Vulpe, 1996). A metal-binding region, represented by the amino acid motif CxxC (where x represents any amino acid) or a methionine- (M-) and/or histidine-(H-) rich sequence, is proposed to be located in the N-terminus of $P_{1B}$-ATPases. The amino acids that flank the invariant proline (P), found approximately 43 amino acids upstream of the DKTGT[LIVM][TIS] motif in all P-type ATPases, vary with substrate specificity (Solioz and Vulpe, 1996). It has been proposed that in $P_{1B}$-ATPases this proline (P) is preceded by a cysteine (C) and followed by a cysteine (C), histidine (H), or serine (S). The term CPx-ATPases has, consequently, been used to describe the $P_{1B}$-ATPases (Solioz and Vulpe, 1996). A HP dipeptide has also been identified in $P_{1B}$-ATPases. This motif is located 34 to 43 amino acids downstream of the DKTGT[LIVM][TIS] motif. The equivalent position in other P-type ATPases does not appear to contain conserved amino acids.

The Type IB subfamily can be divided into at least two groups, the $Cu^{2+}$- and $Cd^{2+}$-ATPases (Axelsen and Palmgren, 1998). The CadA cadmium resistance determinant from a *Staphylococcus aureus* plasmid was the first $P_{1B}$-ATPase amino acid sequence described (Nucifora *et al.*, 1989; Silver *et al.*, 1989). The CadA protein mediates ATP-dependent, $Cd^{2+}$-efflux (Tsai *et al.*, 1992) that involves a phosphorylated intermediate (Tsai and Linet, 1993). Two $P_{1B}$-ATPases were later discovered in *Enterococcus hirae* as a result of a cross-reacting, $P_{1A}$-ATPase antibody (Odermatt *et al.*, 1992). The genes encoding these $P_{1B}$-ATPases were named *CopA* and *CopB*, since disruption of the genes rendered the cells copper-dependent and copper-sensitive, respectively. These P-type ATPases were, consequently, proposed to transport copper (Odermatt *et al.*, 1993) and represented the first members of the $Cu^{2+}$-ATPase group. A subsequent study confirmed that CopB does transport copper and also suggested that silver may be transported (Solioz and Odermatt, 1995).

The $P_{1B}$-ATPases were originally thought to be limited to bacterial species (Solioz and Vulpe, 1996). The cloning of two $P_{1B}$-ATPases responsible for the human Menkes disease (*ATP7A*; Chelly *et al.*, 1993; Mercer *et al.*, 1993; Silver *et al.*, 1993; Vulpe *et al.*, 1993) and Wilson disease (*ATP7B*; Bull *et al.*, 1993; Tanzi *et al.*, 1993) revealed a more widespread presence and a more general role for the $P_{1B}$-ATPases. The identification of ATP7A and ATP7B has resulted in a large body of literature focusing on these $P_{1B}$-ATPases. As an example, studies have focused on mutational analysis of conserved amino acids, determining the function of the metal-binding domains, and examining the intracellular location(s) of the proteins (Hung *et al.*, 1997; Lutsenko *et al.*, 1997; Iida *et al.*, 1998; Strausak *et al.*, 1999; Suzuki and Gitlin, 1999; Forbes and Cox 1998, 2000).

Two $P_{1B}$-ATPases are encoded by the genome of the model eukaryote *Saccharomyces cerevisiae* (Catty *et al.*, 1997). The first $P_{1B}$-ATPase gene, *PCA1* (Rad *et al.*, 1994), has been proposed to encode a $Cu^{2+}$-ATPase. The evidence supporting a role for the Pca1 protein in copper transport is limited and the function of the *PCA1* gene product is not known (Rad *et al.*, 1994). The second $P_{1B}$-ATPase gene, *CCC2* (Fu *et al.*, 1995), has been shown to encode a $Cu^{2+}$-ATPase that plays an important role in establishing a functional high-affinity iron uptake system (Yuan *et al.*, 1995). Yeast strains in which the *CCC2* gene has been disrupted have become an important tool for studying other putative $Cu^{2+}$-ATPases. The ability of a putative $Cu^{2+}$-ATPase to rescue a *ccc2* mutant has been interpreted as evidence suggesting that the protein does transport copper (Hung *et al.*, 1997; Sambongi *et al.*, 1997; Forbes and Cox; 1998; Iida *et al.*, 1998; Payne and Gitlin, 1998; Forbes, 2000). An understanding of how the Ccc2 protein (Ccc2p) functions in *S. cerevisiae* cells is, therefore, important. Figure 1.2 provides a brief overview of the role that Ccc2p plays in copper and iron homeostasis of *S. cerevisiae*.

## 1.5 The *Saccharomyces cerevisiae CCC2* gene encodes a $Cu^{2+}$-ATPase that plays an important role in copper and iron homeostasis

A great deal of research has focused on elucidating the function of the Ccc2p in *S. cerevisiae*. The Ccc2p amino acid sequence shares the greatest sequence identity with $P_{1B}$-ATPases and was grouped with the $Cu^{2+}$-ATPases by Axelsen and Palmgren (1998). The *CCC2* gene product has been shown to play an important role in intracellular copper transport and the production of a functional high-affinity iron uptake system (Yuan *et al.*, 1995, 1997). The Ccc2 protein is localized to a late- or post-Golgi compartment in the secretory pathway (Yuan *et al.*, 1997), where it delivers copper to the multicopper oxidase encoded by the *FET3* gene (Yuan *et al.*, 1995, 1997). The copper loaded Fet3 protein (holoFet3p) is required at the plasma membrane for high-affinity iron uptake (Askwith *et al.*, 1994). *Saccharomyces cerevisiae* strains that have had the *CCC2* gene disrupted produce copper-deficient Fet3p (apoFet3p) and are consequently deficient in high-affinity iron uptake (Yuan *et al.*, 1995).

## 1.6 Specific components of the copper homeostasis pathway are required for the delivery of copper to the Ccc2 and Fet3 proteins

The delivery of copper to Ccc2p is dependent upon copper uptake across the plasma membrane. Under copper-limited conditions, copper enters the cell by a high-affinity uptake system. This system requires that $Cu^{2+}$ first be reduced to $Cu^{+}$ (Hassett *et al.*, 1995). This reduction is mediated by the plasma membrane reductases encoded by the *FRE1* and *FRE2* genes (Georgatsou *et al.*, 1997). Once reduced, $Cu^{+}$ can be transported across the plasma membrane by the high-affinity copper transport proteins encoded by the *CTR1* (Dancis *et al.*, 1994b) and *CTR3* (Knight *et al.*, 1996) genes. A deficiency in either of these proteins alone does not result in a loss of high-affinity copper

uptake. High-affinity copper uptake is, however, absent when both the Ctr1 and Ctr3 proteins are inactive (Knight *et al.*, 1996). Expression of the *CTR3* gene in most *S. cerevisiae* laboratory strains is eliminated by a Ty transposon insertion in the *CTR3* promoter region and high-affinity copper uptake is dependent solely on the Ctr1 protein (Ctr1p; Knight *et al.*, 1996). Copper uptake under copper-replete conditions occurs via a low-affinity uptake system (Dancis *et al.*, 1994b) that remains to be fully characterized. A loss of copper uptake results in a global copper deficiency in the cell and a number of copper-starvation related phenotypes (Dancis *et al.*, 1994). The loss of functional, copper-containing Fet3p and high-affinity iron uptake may, therefore, only account for a portion of the phenotype observed when *S. cerevisiae* cells are deficient in copper uptake.

The concentration of free copper available in the cytoplasm of *S. cerevisiae* has been estimated to be less than $10^{-18}$ M, which is equivalent to many orders of magnitude less than one atom of free copper per cell (Rae *et al.*, 1999). Small, cytosolic proteins, known as metallochaperones, help to regulate the activity of copper in the cytoplasm (Glerum *et al.*, 1996; Culotta *et al.*, 1997; Lin *et al.*, 1997; Rae *et al.*, 1999). Three copper metallochaperones have been identified in *S. cerevisiae* and each of these proteins, encoded by the *ATX1*, *COX17*, and *LYS7* genes, delivers copper to a specific intracellular target. The Lys7 (Culotta *et al.*, 1997), Cox17 (Glerum *et al.*, 1996), and Atx1 (Lin *et al.*, 1997) proteins deliver copper to the cytosolic copper/zinc superoxide dismutase, the mitochondria, and the Ccc2 protein, respectively. The Atx1 protein (Atx1p), therefore, participates in the formation of holoFet3p and the establishment of a functional high-affinity iron uptake system (Lin *et al.*, 1997; Klomp *et al.*, 1997). A partial inhibition of high-affinity iron uptake is observed when the *ATX1* gene is disrupted, suggesting that Ccc2p and Fet3p can also obtain copper in an *ATX1*-independent manner (Lin *et al.*, 1997). A mechanism involving endocytosis is thought to mediate this *ATX1*-independent delivery of copper to Ccc2p and Fet3p (Lin *et al.*, 1997).

## 1.7 The Fet3, Ftr1, Fre1, and Fre2 proteins are required for high-affinity iron uptake

The high-affinity iron uptake system is dependent upon a number of proteins in addition to Fet3p. Iron is commonly found in nature as the sparingly soluble $Fe(OH)_3$ and organisms, consequently, require mechanisms to solubilize the $Fe^{3+}$. *Saccharomyces cerevisiae* reduces extracellular $Fe^{3+}$ to the more soluble $Fe^{2+}$ using reductases encoded by the *FRE1* (Dancis *et al.*, 1992) and *FRE2* (Georgatsou *et al.*, 1994) genes. These are the same reductases that are involved in producing the $Cu^{+}$ that is required for high-affinity copper uptake (Georgatsou *et al.*, 1997). The Fet3 protein shares a high level of similarity with a family of multicopper oxidase proteins (Askwith *et al.*, 1994) and has been shown to exhibit a copper-dependent ferroxidase activity *in vitro* (Yuan *et al.*, 1995; de Silva *et al.*, 1997). One model of the high-affinity iron uptake system (Eide, 1998) suggests that Fet3p oxidizes the $Fe^{2+}$ that is produced by the reductases to $Fe^{3+}$. This $Fe^{3+}$ is then transferred to a $Fe^{3+}$-binding site on the protein encoded by the *FTR1* gene. The Ftr1 protein (Ftr1p) is thought to act as a permease that undergoes a conformational change and releases the $Fe^{3+}$ to the cytoplasm. The possibility that Fet3p oxidizes Ftr1p instead of $Fe^{2+}$ has not been formally excluded (Radisky and Kaplan, 1999). The Fet3 and Ftr1 proteins have been suggested to function as a complex (Stearman *et al.*, 1996). This suggestion is based upon the observation that Fet3p is required for Ftr1p to be transported to the plasma membrane and that Ftr1p is necessary for the maturation of Fet3p through the secretory pathway. Copper loading of Fet3p is not required for localization of Fet3p and Ftr1p to the plasma membrane. *Saccharomyces cerevisiae* strains that are unable to deliver copper to Fet3p, consequently, contain inactive, copper-deficient Fet3p (apoFet3p) at the cell surface.

## 1.8 Regulation of high-affinity iron and copper uptake

The high-affinity iron uptake system is induced under iron-limited conditions and repressed under iron-replete conditions (Dancis *et al.*, 1992; Askwith *et al.*, 1994;

Yamaguchi-Iwai *et al.*, 1995, 1996; Lin *et al.*, 1997). The transcription factor encoded by the *AFT1* gene plays an important role in the regulation of high-affinity iron uptake (Yamaguchi-Iwai *et al.*, 1995, 1996; Lin *et al.*, 1997). The Aft1 protein (Aft1p) regulates the *FTR1*, *FET3*, *FRE1*, *FRE2*, *CCC2*, and *ATX1* genes. Some of these genes, such as *FTR1* and *FET3*, encode proteins that are directly involved in high-affinity iron uptake. Other genes, including *CCC2* and *ATX1*, encode components of the copper homeostasis pathway that are involved in the delivery of copper to Fet3p. The remaining genes, *FRE1* and *FRE2*, are important for both high-affinity iron uptake and copper delivery to Fet3p. The Aft1 protein acts as a transcriptional activator that binds DNA when iron levels are low, but not when iron levels are high (Yamaguchi-Iwai *et al.*, 1995, 1996). It has been proposed that iron binds Aft1p under iron-replete conditions and that this binding precludes DNA binding and transcriptional activation.

High-affinity copper uptake is repressed when copper levels are high and induced when copper levels are low (Dancis *et al.*, 1994, Knight *et al.*, 1996; Labbé *et al.*, 1997; Yamaguchi-Iwai *et al.*, 1997). The *MAC1* gene encodes a transcription factor that mediates the expression of the *CTR1*, *CTR3*, and *FRE1* genes whose products are involved in high-affinity copper uptake (Labbé *et al.*, 1997; Yamaguchi-Iwai *et al.*, 1997). The Mac1 protein (Mac1p) is able to bind to specific DNA sequences and induce transcription of the corresponding genes. This DNA binding and transcriptional activation does not occur when copper levels are high; Mac1p binds copper instead of DNA under these conditions (Jensen and Winge, 1998). In addition to preventing Mac1p from binding DNA, high concentrations of copper actually result in a rapid degradation of Mac1p (Zhu *et al.*, 1998). Levels of Ctr1p are also regulated by a copper-dependent proteolysis (Ooi *et al.*, 1996).

## 1.9 Low-affinity transport systems function under metal-replete conditions

Both high- and low-affinity uptake systems mediate the uptake of metals across the *S. cerevisiae* plasma membrane (Eide, 1998). High-affinity uptake systems, which are usually specific for a single metal, function when the metal levels are low. Low-affinity uptake systems mediate metal uptake under metal-replete conditions and may participate in the transport of a number of different substrates. The *FET4* gene encodes a low-affinity iron uptake system that is specific for $Fe^{2+}$ over $Fe^{3+}$(Dix *et al.*, 1994). The Fet4 protein (Fet4p) may also participate in the transport of a number of different substrates. High concentrations of $Ni^{2+}$, $Cd^{2+}$, $Co^{2+}$, and $Cu^{2+}$ inhibit Fet4p-mediated $Fe^{2+}$ uptake (Dix *et al.*, 1994) and overexpression of the *FET4* gene results in a hypersensitivity to these metals (Dix, 1996 as reported in Eide 1998). *Saccharomyces cerevisiae* strains that have the *FET3* gene disrupted are deficient in high-affinity iron uptake and display an increased metal sensitivity to $Zn^{2+}$, $Mn^{2+}$, $Co^{2+}$, and $Cu^{2+}$ (Li and Kaplan, 1998). This increased sensitivity is thought to be due to an increase in uptake of the metals as a result of increased expression of low-affinity systems such as the one encoded by *FET4*. A low-affinity copper uptake system has also been reported (Dancis *et al.*, 1994b) and a recent study suggests that Fet4p may mediate this low-affinity copper uptake (Hassett *et al.*, 2000).

## 1.10 Metal supplementation can rescue growth defects that are due to a defective high-affinity iron uptake system

A particular pattern of metal supplementation has been found to rescue *S. cerevisiae* stains that display growth defects that result from a deficiency in high-affinity iron uptake (Yuan *et al.*, 1995). The literature describes two major growth defects that result from the loss of high-affinity iron uptake. The first defect is an inability to grow on

non-fermentable media that is thought to be due to the absence of cofactor(s) for cytochrome-*c* oxidase (Askwith *et al.*, 1994; Yuan *et al.*, 1995). The inability to grow on iron-limited media is the second defect and results from a global iron-deficiency (Askwith *et al.*, 1994; Forbes and Cox, 1998). *Saccharomyces cerevisiae* strains in which a component of the high-affinity iron uptake system is inactive can be rescued by supplementing the media with iron. This iron-mediated rescue is believed to result from the ability of low-affinity uptake systems to provide the necessary iron under iron-replete conditions. Iron-mediated growth rescue has been demonstrated for *fet3* mutants on non-fermentable (Yuan *et al.*, 1995) and iron-limited media (Forbes and Cox, 1998). When high-affinity iron uptake deficiency is a result of the inability to load Fet3p with copper a similar iron-mediated rescue may be possible.

The major defect apparent in *ccc2* mutants is an inability to deliver copper to Fet3p and a resulting loss of high-affinity iron uptake (Yuan *et al.*, 1995). The *ccc2* mutants, therefore, exhibit an iron-mediated rescue on non-fermentable (Yuan *et al.*, 1995) and iron-limited media (Forbes and Cox, 1998) that is similar to that described for the *fet3* mutants. While the *ccc2* mutants display a copper deficiency that appears to be limited to a single protein (Yuan *et al.*, 1995), the *ctr1* mutants are globally copper deficient (Dancis *et al.*, 1994, 1994b). Since both copper and iron are required as cofactors for cytochrome-*c* oxidase, iron-supplementation is unable to rescue the growth of *ctr1* mutants on non-fermentable media (Yuan *et al.*, 1995). Iron-mediated rescue may be possible for *ctr1* mutants on iron-limited media, since this growth defect is more intimately associated with the deficiency in high-affinity iron uptake. When the high-affinity iron uptake deficiency is a result of the inability to load Fet3p with copper, a copper-mediated rescue is possible. Copper loading of apoFet3p (Fet3p protein that has not been loaded with copper) at the plasma membrane has been proposed to occur when the concentration of copper in the medium exceeds 1 μM (Yuan *et al.*, 1997). This proposal is based on the observation that the concentration-dependent, copper-mediated rescue of high-affinity iron uptake was found to be similar in *ccc2* single mutants and *ccc2ctr1* double mutants. This suggests that the ability of *ccc2* mutants to grow on media that has been supplemented with copper results from mechanisms other than copper

uptake across the plasma membrane (Yuan *et al.*, 1997). The observation that copper is able to restore high-affinity iron uptake in copper-deficient cells independent of *de novo* protein synthesis (Askwith *et al.*, 1994) also suggests that existing proteins are being loaded with copper. The *ctr1* mutants are globally copper-deficient and while copper loading of apoFet3p could occur at the plasma membrane, the remaining copper deficient proteins, which are not located at the cell surface, would acquire copper as a result of a low affinity uptake system. Copper-mediated growth rescue has been demonstrated for *ccc2* and *ctr1* mutants on non-fermentable (Yuan *et al.*, 1995) and iron-limited media (Forbes and Cox, 1998).

## 1.11 *Saccharomyces cerevisiae* strains other than the *ccc2* mutant may be of use for studying $P_{1B}$-ATPases

Complementation of the *ccc2* mutant by a putative $P_{1B}$-ATPase suggests that the putative $P_{1B}$-ATPase is a $Cu^{2+}$-ATPase that localizes to the secretory system. The $P_{1B}$-ATPases that have been characterized in bacteria (Section 1.4) suggest that $Cd^{2+}$- and $Cu^{2+}$-ATPases might exist in the plasma membrane of eukaryotic cells. Other *S. cerevisiae* strains may aid in the characterization of such $P_{1B}$-ATPases. *Saccharomyces cerevisiae* strains that are deficient in high-affinity copper uptake, due to a mutated *CTR1* gene, have been used to identify eukaryotic genes that mediate copper uptake (Kampfenkel *et al.*, 1995; Zhou *et al.*, 1997). The eukaryotic proteins identified were not P-type ATPases. A similar study in which a putative $Cu^{2+}$-ATPase is able to rescue a *ctr1* mutant could suggest that the putative $P_{1B}$-ATPase transports copper into the cell. Most *S. cerevisiae* strains are sensitive to low concentrations of cadmium; there are, however, some strains that are resistant to higher levels of cadmium (Tohoyama *et al.*, 1990). The relative sensitivity/resistance of *S. cerevisiae* strains to cadmium can be examined by observing growth on cadmium concentration-gradient plates (Tohoyama *et al.*, 1990). If a *S. cerevisiae* strain expressing a putative $P_{1B}$-ATPase exhibits an altered growth pattern in the presence of cadmium it is likely that the $P_{1B}$-ATPase mediates cadmium transport.

## 1.12 $P_{1B}$-ATPases in plants?

Axelsen and Palmgren (1998) searched the European Molecular Biology Laboratory (EMBL; release 51) and SWISS-PROT (release 34) databases in an attempt to identify all full-length, putative P-type ATPases. Although they did identify some P-type ATPases from plant species, they did not identify any plant $P_{1B}$-ATPases. Axelsen and Palmgren (1998) did find expressed sequence tag (EST) data suggesting that $P_{1B}$-ATPases exist in the model plant *Arabidopsis thaliana*. Since P-type ATPases that transport the same substrate tend to cluster into families, subfamilies, and groups (Axelsen and Palmgren, 1998) and the $P_{1B}$-ATPase subfamily contains $Cu^{2+}$-ATPase and $Cd^{2+}$-ATPase groups, $P_{1B}$-ATPases that are identified in plants might be involved in copper or cadmium transport.

Copper is an essential micronutrient for plant metabolism. Copper is able to cycle between a stable oxidised ($Cu^{2+}$) and unstable reduced ($Cu^{+}$) state and, consequently, acts as an effective electron acceptor and donor in the active sites of many proteins involved in oxidation and reduction reactions. These copper-containing proteins play fundamental roles in the survival of plants. Copper is a cofactor in proteins required for the electron transfer reactions of respiration (cytochrome-*c* oxidase, alternate oxidase) and photosynthesis (plastocyanin), the detoxification of superoxide radicals (copper/zinc superoxide dismutase), and lignification of plant cell walls (laccase). While copper is required in limited amounts, it is toxic in excess. This toxicity is due to the ability of copper to bind to and inactivate proteins, to interfere with the homeostasis of other metals, and to generate reactive oxygen intermediates. Because copper is both essential and toxic, mechanisms are required to control the amount and location of copper within cells. Transport of copper across biological membranes plays an important role in this copper homeostasis. The $P_{1B}$-ATPase Ccc2p has been shown to play an important role in copper homeostasis in the model eukaryote *S. cerevisiae* (Yuan *et al.*, 1995). Similar roles for $P_{1B}$-ATPases have also been reported in other eukaryotes (Bull *et al.*, 1993; Chelly *et al.*, 1993; Mercer *et al.*, 1993; Silver *et al.*, 1993; Vulpe *et al.*, 1993; Tanzi *et al.*, 1993). Plant $P_{1B}$-ATPases, if they exist, can also be predicted to participate in copper

homeostasis. A functional homologue of the *S. cerevisiae* copper metallochaperone Atx1p, that is required to deliver copper to Ccc2p, has been identified in *Arabidopsis* (Himelblau *et al.*, 1998). The identification of this metallochaperone, named Copper CHaperone (CCH), suggests that plants may also contain a Ccc2p functional homologue.

Cadmium is a non-essential, toxic metal that has been designated a human carcinogen (Waalkes, 2000). The toxic effects of cadmium are proposed to result from its interference with zinc- and iron-mediated processes. It has been suggested that $Cd^{2+}$ inactivates proteins by displacing the $Zn^{2+}$ and $Fe^{2+}$ cofactors (Stohs and Bagchi, 1995) and that the displaced iron may generate reactive hydroxyl radicals (Vido *et al.*, 2001). Cadmium has, indeed, been found to induce oxidative stress (Brennan and Schiestl, 1996). The mechanism of cadmium carcinogenesis remains to be determined, but may involve altered zinc metabolism, cell proliferation, or DNA repair (Waalkes, 2000). Food is the major source of cadmium exposure in nonsmokers and the uptake of cadmium by agricultural plants is a major cause for the accumulation of cadmium in the human body (Wagner, 1993; Van Bruwaene *et al.*, 1984). Understanding cadmium transport may aid in developing plants that accumulate lower levels of cadmium. Alternatively, understanding cadmium transport may lead to the development of plants that hyperaccumulate cadmium and can be used in phytoremediation. In prokaryotes, $P_{1B}$-ATPases clearly participate in cadmium efflux (Nucifora *et al.*, 1989; Silver *et al.*, 1989; Tsai *et al.*, 1992) and it is possible that $P_{1B}$-ATPases may also participate in cadmium transport in eukaryotes including plants.

The possibility that $P_{1B}$-ATPases may be found in plants is an exciting proposition. When a $P_{1B}$-ATPase is identified, it is often possible to predict its substrate specificity based on an analysis of the amino acid sequence. The $P_{1B}$-ATPase subfamily can be divided into at least two groups, the $Cu^{2+}$- and $Cd^{2+}$-ATPases. Plant $P_{1B}$-ATPases might, therefore, be involved in copper or cadmium transport and, as discussed above, there are important ramifications that could result from a clearer understanding of the transport of copper and cadmium in plant cells. The $P_{1B}$-ATPases are also of interest since they represent relatively simple systems to study as individual proteins appear to

transport a specific substrate and depend on a single subunit for their activity. There is also a large body of literature that, due to the similarities of the P-type ATPases, can be utilized to make general predictions about the structure and catalytic mechanisms of newly identified P-type ATPases.

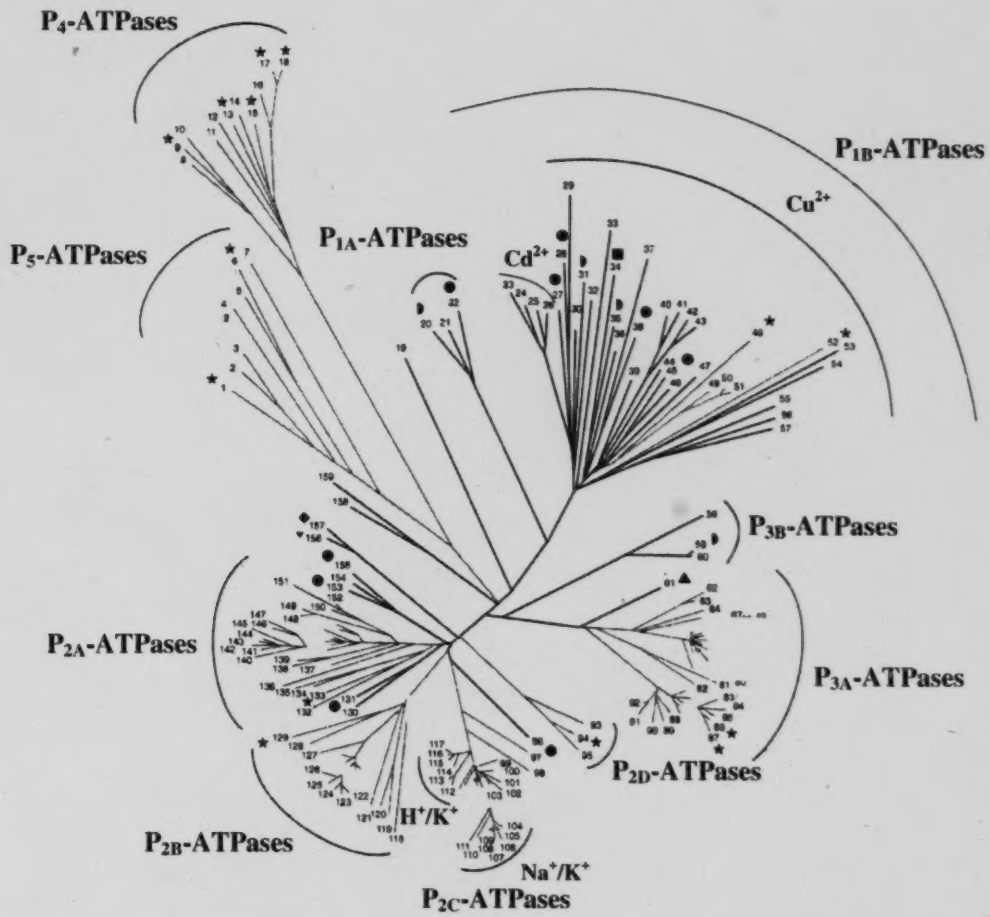


**Figure 1.1 The P-type ATPase superfamily.** The P-type ATPase superfamily can be divided into five major families ($P_1$-, $P_2$-, $P_3$-, $P_4$-, and $P_5$-ATPases). The $P_1$-, $P_2$-, and $P_3$-ATPases can be further divided into $P_{1A}$- and $P_{1B}$-ATPase subfamilies; $P_{2A}$-, $P_{2B}$-, $P_{2C}$-, and $P_{2D}$-ATPase subfamilies; and $P_{3A}$- and $P_{3B}$-ATPase subfamilies, respectively. The $P_{1B}$-ATPase subfamily contains at least two groups, the $Cd^{2+}$- and $Cu^{2+}$-ATPases. The $P_{2C}$-ATPase subfamily can be divided into $Na^+/K^+$- and $H^+/K^+$-ATPase groups. This figure was adapted from Axelsen and Palmgren (1998).

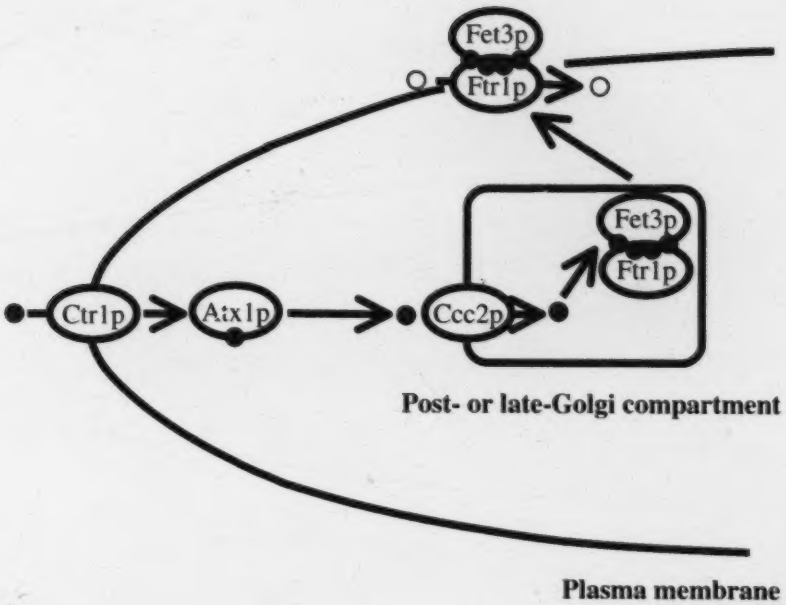


**Figure 1.2 Model illustrating the role of Ccc2p in copper and iron homeostasis of *S. cerevisiae*.** Ccc2p is located in the membrane of a late- or post-Golgi compartment. Ccc2p transports copper (shown as shaded circles) into this compartment where the copper is incorporated into the multicopper oxidase Fet3p. The iron permease Ftr1p is necessary for the maturation of Fet3p through the secretory pathway. The Ftr1 protein and the multicopper Fet3p are then localized to the plasma membrane where they are involved in high-affinity iron uptake (iron is shown as open circles). Ctr1p and Atx1p play important roles in the delivery of copper to Ccc2p. Ctr1p is a plasma membrane protein involved in high-affinity copper uptake. Atx1p is a copper chaperone that carries copper through the cytoplasm and delivers it to Ccc2p.

## 1.13 Literature cited

**Askwith C, Eide D, Van Ho A, Bernard PS, Li L, Davis-Kaplan S, Sipe DM, Kaplan J** (1994) The *FET3* gene of *Saccharomyces cerevisiae* encodes a multicopper oxidase required for ferrous iron uptake. Cell **76(2)**:403-410

**Axelsen KB, Palmgren MG** (1998) Evolution of substrate specificities in the P-type ATPase superfamily. J Mol Evol **46(1)**:84-101

**Brennan RJ, Schiestl RH** (1996) Cadmium is an inducer of oxidative stress in yeast. Mutat Res **356(2)**:171-178

**Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW** (1993) The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. Nat Genet **5(4)**:327-337

**Catty P, de Kerchove d'Exaerde A, Goffeau A** (1997) The complete inventory of the yeast *Saccharomyces cerevisiae* P-type transport ATPases. FEBS Lett **409(3)**:325-332

**Chelly J, Tumer Z, Tonnesen T, Petterson A, Ishikawa-Brush Y, Tommerup N, Horn N, Monaco AP** (1993) Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. Nat Genet **3(1)**:14-19

**Culotta VC, Klomp LW, Strain J, Casareno RL, Krems B, Gitlin JD** (1997) The copper chaperone for superoxide dismutase. J Biol Chem **272(38)**:23469-23472

**Dancis A, Haile D, Yuan DS, Klausner RD** (1994) The *Saccharomyces cerevisiae* copper transport protein (Ctr1p). Biochemical characterization, regulation by copper, and physiologic role in copper uptake. J Biol Chem **269(41)**:25660-25667

**Dancis A, Roman DG, Anderson GJ, Hinnebusch AG, Klausner RD** (1992) Ferric reductase of *Saccharomyces cerevisiae*: molecular characterization, role in iron uptake, and transcriptional control by iron. Proc Natl Acad Sci U S A **89(9)**:3869-3873

**Dancis A, Yuan DS, Haile D, Askwith C, Eide D, Moehle C, Kaplan J, Klausner RD** (1994b) Molecular characterization of a copper transport protein in *Saccharomyces cerevisiae*: an unexpected role for copper in iron transport. Cell **76(2)**:393-402

**de Silva D, Davis-Kaplan S, Fergestad J, Kaplan J** (1997) Purification and characterization of Fet3 protein, a yeast homologue of ceruloplasmin. J Biol Chem **272(22)**:14208-14213

**Dix DR** (1996) Structure, function, and mechanisms of Fet4p, a low affinity transport protein in the yeast *Saccharomyces cerevisiae*. MS thesis. University of Missouri-Columbia, 1996 85 pp

**Dix DR, Bridgham JT, Broderius MA, Byersdorfer CA, Eide DJ** (1994) The *FET4* gene encodes the low affinity Fe(II) transport protein of *Saccharomyces cerevisiae*. J Biol Chem **269(42)**:26092-26099

**Eide DJ** (1998) The molecular biology of metal ion transport in *Saccharomyces cerevisiae*. Annu Rev Nutr **18**:441-469

**Fagan MJ, Saier MH Jr.** (1994) P-type ATPases of eukaryotes and bacteria: sequence analyses and construction of phylogenetic trees. J Mol Evol **38(1)**:57-99

**Forbes JR** (2000) Functional analysis of the copper-transporting P-type ATPase, ATP7B, defective in Wilson disease. PhD thesis. University of Alberta, Canada, 2000 159 pp

**Forbes JR, Cox DW** (1998) Functional characterization of missense mutations in *ATP7B*: Wilson disease mutation or normal variant? Am J Hum Genet **63(6)**:1663-1674

**Forbes JR, Cox DW** (2000) Copper-dependent trafficking of Wilson disease mutant ATP7B proteins. Hum Mol Genet **9(13)**:1927-1935

**Fu D, Beeler TJ, Dunn TM** (1995) Sequence, mapping and disruption of *CCC2*, a gene that cross-complements the Ca(2+)-sensitive phenotype of *csg1* mutants and encodes a P-type ATPase belonging to the Cu(2+)-ATPase subfamily. Yeast **11(3)**:283-292

**Georgatsou E, Alexandraki D** (1994) Two distinctly regulated genes are required for ferric reduction, the first step of iron uptake in *Saccharomyces cerevisiae*. Mol Cell Biol **14(5)**:3065-3073.

**Georgatsou E, Mavrogiannis LA, Fragiadakis GS, Alexandraki D** (1997) The yeast Fre1p/Fre2p cupric reductases facilitate copper uptake and are regulated by the copper-modulated Mac1p activator. J Biol Chem **272(21)**:13786-13792

**Glerum DM, Shtanko A, Tzagoloff A** (1996) Characterization of *COX17*, a yeast gene involved in copper metabolism and assembly of cytochrome oxidase. J Biol Chem **271(24)**:14504-14509

**Hassett R, Dix DR, Eide DJ, Kosman DJ** (2000) The Fe(II) permease Fet4p functions as a low affinity copper transporter and supports normal copper trafficking in *Saccharomyces cerevisiae*. Biochem J **351 Pt 2**:477-484

**Hassett R, Kosman DJ** (1995) Evidence for Cu(II) reduction as a component of copper uptake by *Saccharomyces cerevisiae*. J Biol Chem **270(1)**:128-134

**Himelblau E, Mira H, Lin SJ, Culotta VC, Penarrubia L, Amasino RM** (1998) Identification of a functional homolog of the yeast copper homeostasis gene *ATX1* from *Arabidopsis*. Plant Physiol **117(4)**:1227-1234

**Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD, Gitlin JD** (1997) Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. J Biol Chem **272(34)**:21461-21466

**Iida M, Terada K, Sambongi Y, Wakabayashi T, Miura N, Koyama K, Futai M, Sugiyama T** (1998) Analysis of functional domains of Wilson disease protein (ATP7B) in *Saccharomyces cerevisiae*. FEBS Lett **428(3)**:281-285

**Jensen LT, Winge DR** (1998) Identification of a copper-induced intramolecular interaction in the transcription factor Mac1 from *Saccharomyces cerevisiae*. EMBO J **17(18)**:5400-5408

**Kampfenkel K, Kushnir S, Babiychuk E, Inze D, Van Montagu M** (1995) Molecular characterization of a putative *Arabidopsis thaliana* copper transporter and its yeast homologue. J Biol Chem **270(47)**:28479-28486

**Klomp LW, Lin SJ, Yuan DS, Klausner RD, Culotta VC, Gitlin JD** (1997) Identification and functional expression of *HAH1*, a novel human gene involved in copper homeostasis. J Biol Chem **272(14)**:9221-9226

**Knight SA, Labbe S, Kwon LF, Kosman DJ, Thiele DJ** (1996) A widespread transposable element masks expression of a yeast copper transport gene. Genes Dev **10(15)**:1917-1929

**Labbé S, Zhu Z, Thiele DJ** (1997) Copper-specific transcriptional repression of yeast genes encoding critical components in the copper transport pathway. J Biol Chem **272(25)**:15951-15958

**Lee AG, East JM** (2001) What the structure of a calcium pump tells us about its mechanism. Biochem J **356(Pt 3)**:665-683

**Li L, Kaplan J** (1998) Defects in the yeast high affinity iron transport system result in increased metal sensitivity because of the increased expression of transporters with a broad transition metal specificity. J Biol Chem **273(35)**:22181-22187.

**Lin SJ, Pufahl RA, Dancis A, O'Halloran TV, Culotta VC** (1997) A role for the *Saccharomyces cerevisiae ATX1* gene in copper trafficking and iron transport. J Biol Chem **272(14)**:9215-9220

**Lutsenko S, Kaplan JH** (1995) Organization of P-type ATPases: significance of structural diversity. Biochemistry **34(48)**:15607-15613

**Lutsenko S, Petrukhin K, Cooper MJ, Gilliam CT, Kaplan JH** (1997) N-terminal domains of human copper-transporting adenosine triphosphatases (the Wilson's and Menkes disease proteins) bind copper selectively *in vivo* and *in vitro* with stoichiometry of one copper per metal-binding repeat. J Biol Chem **272(30)**:18939-18944

**MacLennan DH, Green NM** (2000) Structural biology. Pumping ions. Nature **405(6787)**:633-634

**MacLennan DH, Rice WJ, Green NM** (1997) The mechanism of $Ca^{2+}$ transport by sarco(endo)plasmic reticulum $Ca^{2+}$-ATPases. J Biol Chem **272(46)**:28815-28818

**Mercer JF, Livingston J, Hall B, Paynter JA, Begy C, Chandrasekharappa S, Lockhart P, Grimes A, Bhave M, Siemieniak D, *et al.*** (1993) Isolation of a partial candidate gene for Menkes disease by positional cloning. Nat Genet **3(1)**:20-25

**Moller JV, Juul B, le Maire M** (1996) Structural organization, ion transport, and energy transduction of P-type ATPases. Biochim Biophys Acta **1286(1)**:1-51

**Nucifora G, Chu L, Misra TK, Silver S** (1989) Cadmium resistance from *Staphylococcus aureus* plasmid pI258 *cadA* gene results from a cadmium-efflux ATPase. Proc Natl Acad Sci U S A **86(10)**:3544-3548

**Odermatt A, Suter H, Krapf R, Solioz M** (1992) An ATPase operon involved in copper resistance by *Enterococcus hirae*. Ann N Y Acad Sci **671**:484-486

**Odermatt A, Suter H, Krapf R, Solioz M** (1993) Primary structure of two P-type ATPases involved in copper homeostasis in *Enterococcus hirae*. J Biol Chem **268(17)**:12775-12779

**Ooi CE, Rabinovich E, Dancis A, Bonifacino JS, Klausner RD** (1996) Copper-dependent degradation of the *Saccharomyces cerevisiae* plasma membrane copper transporter Ctr1p in the apparent absence of endocytosis. EMBO J **15(14)**:3515-3523

**Palmgren MG, Axelsen KB** (1998) Evolution of P-type ATPases. Biochim Biophys Acta **1365(1-2)**:37-45

**Payne AS, Gitlin JD** (1998) Functional expression of the Menkes disease protein reveals common biochemical mechanisms among the copper-transporting P-type ATPases. J Biol Chem **273(6)**:3765-3770

**Rad MR, Kirchrath L, Hollenberg CP** (1994) A putative P-type Cu(2+)-transporting ATPase gene on chromosome II of *Saccharomyces cerevisiae*. Yeast **10(9)**:1217-1225

**Radisky D, Kaplan J** (1999) Regulation of transition metal transport across the yeast plasma membrane. J Biol Chem **274(8)**:4481-4484

**Rae TD, Schmidt PJ, Pufahl RA, Culotta VC, O'Halloran TV** (1999) Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. Science **284(5415)**:805-808

**Sambongi Y, Wakabayashi T, Yoshimizu T, Omote H, Oka T, Futai M** (1997) *Caenorhabditis elegans* cDNA for a Menkes/Wilson disease gene homologue and its function in a yeast *CCC2* gene deletion mutant. J Biochem (Tokyo) **121(6)**:1169-1175

**Silver S, Nucifora G, Chu L, Misra TK** (1989) Bacterial resistance ATPases: primary pumps for exporting toxic cations and anions. Trends Biochem Sci **14(2)**:76-80

**Silver S, Nucifora G, Phung LT** (1993) Human Menkes X-chromosome disease and the *Staphylococcal* cadmium-resistance ATPase: a remarkable similarity in protein sequences. Mol Microbiol **10(1)**:7-12

**Solioz M, Odermatt A** (1995) Copper and silver transport by CopB-ATPase in membrane vesicles of Enterococcus hirae. J Biol Chem **270(16)**:9217-9221

**Solioz M, Vulpe C** (1996) CPx-type ATPases: a class of P-type ATPases that pump heavy metals. Trends Biochem Sci **21(7)**:237-241

**Stearman R, Yuan DS, Yamaguchi-Iwai Y, Klausner RD, Dancis A** (1996) A permease-oxidase complex involved in high-affinity iron uptake in yeast. Science **271(5255)**:1552-1557

**Stohs SJ, Bagchi D** (1995) Oxidative mechanisms in the toxicity of metal ions. Free Radic Biol Med **18(2)**:321-336

**Strausak D, La Fontaine S, Hill J, Firth SD, Lockhart PJ, Mercer JF** (1999) The role of GMXCXXC metal binding sites in the copper-induced redistribution of the Menkes protein. J Biol Chem **274(16)**:11170-11177

**Suzuki M, Gitlin JD** (1999) Intracellular localization of the Menkes and Wilson's disease proteins and their role in intracellular copper transport. Pediatr Int **41(4)**:436-442

**Tanzi RE, Petrukhin K, Chernov I, Pellequer JL, Wasco W, Ross B, Romano DM, Parano E, Pavone L, Brzustowicz LM, *et al.*** (1993) The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. Nat Genet **5(4)**:344-350

**Tohoyama H, Inouhe M, Joho M, Murayama T** (1990) Resistance to cadmium is under control of the *CAD2* gene in the yeast *Saccharomyces cerevisiae*. Curr Genet **18(3)**:181-185

**Toyoshima C, Nakasako M, Nomura H, Ogawa H** (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. Nature **405(6787)**:647-655

**Toyoshima C, Sasabe H, Stokes DL** (1993) Three-dimensional cryo-electron microscopy of the calcium ion pump in the sarcoplasmic reticulum membrane. Nature **362(6419)**:467-471

**Tsai KJ, Linet AL** (1993) Formation of a phosphorylated enzyme intermediate by the *cadA* Cd(2+)-ATPase. Arch Biochem Biophys **305(2)**:267-270

**Tsai KJ, Yoon KP, Lynn AR** (1992) ATP-dependent cadmium transport by the *cadA* cadmium resistance determinant in everted membrane vesicles of *Bacillus subtilis*. J Bacteriol **174(1)**:116-121

**Van Bruwaene R, Kirchmann R, Impens R** (1984) Cadmium contamination in agriculture and zootechnology. Experientia **40(1)**:43-52

**Vido K, Spector D, Lagniel G, Lopez S, Toledano MB, Labarre J** (2001) A proteome analysis of the cadmium response in *Saccharomyces cerevisiae*. J Biol Chem **276(11)**:8469-8474

**Voskoboinik I, Mar J, Strausak D, Camakaris J** (2001) The regulation of catalytic activity of the Menkes copper-translocating P-type ATPase. Role of high affinity copper-binding sites. J Biol Chem **276(30)**:28620-28627

**Vulpe C, Levinson B, Whitney S, Packman S, Gitschier J** (1993) Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nat Genet **3(1)**:7-13

**Waalkes MP** (2000) Cadmium carcinogenesis in review. J Inorg Biochem **79(1-4)**:241-244

**Wagner GJ** (1993) Accumulation of cadmium in crop plant and its consequences to human health. Advances in Agronomy **51**:173-212

**Yamaguchi-Iwai Y, Dancis A, Klausner RD** (1995) AFT1: a mediator of iron regulated transcriptional control in *Saccharomyces cerevisiae*. EMBO J **14(6)**:1231-1239

**Yamaguchi-Iwai Y, Stearman R, Dancis A, Klausner RD** (1996) Iron-regulated DNA binding by the AFT1 protein controls the iron regulon in yeast. EMBO J **15(13)**:3377-3384

**Yamaguchi-Iwai Y, Serpe M, Haile D, Yang W, Kosman DJ, Klausner RD, Dancis A** (1997) Homeostatic regulation of copper uptake in yeast via direct binding of MAC1 protein to upstream regulatory sequences of *FRE1* and *CTR1*. J Biol Chem **272(28)**:17711-17718

**Yuan DS, Dancis A, Klausner RD** (1997) Restriction of copper export in *Saccharomyces cerevisiae* to a late Golgi or post-Golgi compartment in the secretory pathway. J Biol Chem **272(41)**:25787-25793

**Yuan DS, Stearman R, Dancis A, Dunn T, Beeler T, Klausner RD** (1995) The Menkes/Wilson disease gene homologue in yeast provides copper to a ceruloplasmin-like oxidase required for iron uptake. Proc Natl Acad Sci U S A **92(7)**:2632-2636

**Zhang P, Toyoshima C, Yonekura K, Green NM, Stokes DL** (1998) Structure of the calcium pump from sarcoplasmic reticulum at 8-A resolution. Nature **392(6678)**:835-839

**Zhou B, Gitschier J** (1997) *hCTR1*: a human gene for copper uptake identified by complementation in yeast. Proc Natl Acad Sci U S A **94(14)**:7481-7486

**Zhu Z, Labbe S, Pena MM, Thiele DJ** (1998) Copper differentially regulates the activity and degradation of yeast Mac1 transcription factor. J Biol Chem **273(3)**:1277-1280

# 2 The identification, analysis, and cloning of plant sequences corresponding to putative $P_{1B}$-ATPases

## 2.1 Introduction

The $P_{1B}$-ATPases, which were first identified as proteins that mediate prokaryotic cadmium resistance (Silver *et al.*, 1989; Nucifora *et al.*, 1989), were initially thought to be limited to bacterial species (Solioz and Vulpe, 1996). The cloning of the genes responsible for the human Menkes disease (*ATP7A*; Chelly *et al.*, 1993; Mercer *et al.*, 1993; Silver *et al.*, 1993; Vulpe *et al.*, 1993) and Wilson disease (*ATP7B*; Bull *et al.*, 1993; Tanzi *et al.*, 1993) revealed that $P_{1B}$-ATPases are also found in eukaryotes. A search of the European Molecular Biology Laboratory (EMBL; release 51) and SWISS-PROT (release 34) databases that was performed by Axelsen and Palmgren (1998) a few years later did not, however, identify any full-length plant $P_{1B}$-ATPases. Expressed sequence tag (EST) data (December, 1996) did, nevertheless, suggest that $P_{1B}$-ATPases exist in the model plant *Arabidopsis* (Axelsen and Palmgren, 1998).

Sequence databases grow at an exponential rate and full-length sequences that encode $P_{1B}$-ATPases from plants may have been added to the databases in the approximately one and a half year period between the time that Axelsen and Palmgren (1998) searched the databases and the time that this project was initiated. The large-scale sequencing efforts coordinated by the *Arabidopsis* Genome Initiative (AGI; Meinke *et al.*, 1998) increase the probability that full-length $P_{1B}$-ATPase sequences from *Arabidopsis* may have been added to the databases.

Expressed sequence tags (ESTs) are an important resource as they provide a glimpse at the transcripts that are being produced by a particular organism. Although ESTs usually only reveal a portion of a full-length coding sequence, the information provided by the EST may be sufficient to speculate on the type of protein that is encoded by the corresponding full-length coding sequence. Axelsen and Palmgren (1998)

identified a number of *Arabidopsis* ESTs that suggested that $P_{1B}$-ATPases exist in this plant. Additional *Arabidopsis* ESTs, which might correspond to other $P_{1B}$-ATPases, may also have been added to the databases since Axelsen and Palmgren (1998) performed their search. Expressed sequence tags (ESTs) suggesting that $P_{1B}$-ATPases exist in other plant species might also be found in the databases. A match can sometimes be made between a full-length genomic sequence and an EST. When such a match is made, the EST provides evidence suggesting that the corresponding gene is being transcribed and confirms regions of the predicted coding and protein sequences.

Although identification of full-length sequences and/or ESTs could provide some evidence to suggest that $P_{1B}$-ATPases exist in plants, a cloned cDNA that encodes a putative $P_{1B}$-ATPase would provide a stronger argument. A cloned, full-length cDNA could confirm the predicted coding sequence of a genomic entry or reveal a novel coding sequence that does not correspond to a previously sequenced gene. A cloned cDNA would also permit additional experiments, such as a complementation assay, that might demonstrate that the putative $P_{1B}$-ATPase is functional and provide information about its substrate specificity. In addition to providing information about plant $P_{1B}$-ATPases that might exist, identification of full-length sequences and/or ESTs from databases might provide information that could be used to develop a cloning strategy to isolate a plant cDNA that encodes a putative $P_{1B}$-ATPase. The objective of this study was, therefore, to identify sequence information that was available for putative $P_{1B}$-ATPases from plants and use this information to develop a strategy to clone plant cDNA(s) that encode putative $P_{1B}$-ATPase(s).

## 2.2 Materials and methods

### 2.2.1 The identification and analysis of plant sequences that correspond to putative $P_{1B}$-ATPases

#### 2.2.1.1 Literature search

Various databases were searched to identify putative or known plant $P_{1B}$-ATPases described in the literature. The Agricola, Biological and Agricultural Index, CAB Abstracts, and Current Contents databases were searched online through the University of Alberta Library (http://www.library.ualberta.ca). The PubMed database was accessed through the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov; Wheeler *et al.*, 2001) and searched using the Entrez retrieval system (Schuler *et al.*, 1996).

#### 2.2.1.2 Sequence database search: Basic Local Alignment Search Tool at the National Center for Biotechnology Information

The Basic Local Alignment Search Tool (BLAST) family of search programs (BLAST; Altschul *et al.*, 1990) that is supported by NCBI was used to identify sequences based on sequence similarities. The PAA1 (BAA23769/2668492), CadA (AAB59154/150719), CopB (AAA61836/290643), and Ccc2p (AAC37425/538515) amino acid sequences were used as queries that were compared against a protein database (blastp) or a nucleotide database dynamically translated in all reading frames (tblastn). All of the Type IB (Section 1.4) and P-type (Section 1.1) motifs are present in these sequences. The PAA1 sequence, which was identified by the literature search (Section 2.3.1.1), is a putative plant $P_{1B}$-ATPase. The CadA (*Staphylococcus aureus*) and CopB (*Enterococcus hirae*) proteins have been shown to participate in $Cd^{2+}$ (Tsai *et al.*, 1992) and $Cu^{2+}$ (Solioz and Odermatt, 1995) transport, respectively. The copper-transporting $P_{1B}$-ATPase Ccc2p is from *Saccharomyces cerevisiae*, a model organism that will be used

in a complementation assay to analyze the transport activity of protein(s) encoded by cDNA(s) isolated in this project.

The blastp and tblastn search parameters were adjusted as follows. The NCBI-gi option was selected so that the GenInfo (GI) identifier would be reported for all matches. Advanced BLAST searches, which offer the option of entering an organism or taxonomic class, were used so that the searches could be limited to the taxonomic classes Viridiplantae (green plants) or Embryophyta (higher plants). When using the blastp program, the default "nr" database was searched. The "nr" database contains all non-redundant protein sequences from Protein Data Bank (PDB; Berman *et al.*, 2000), SwissProt (Bairoch and Apweiler, 2000), International Protein Sequence Database (PIR; Barker *et al.*, 2001), Protein Research Foundation (PRF), and translations from annotated coding regions in GenBank (Benson *et al.*, 2000), EMBL (Stoesser *et al.*, 2001), and the DNA Database of Japan (DDBJ; Tateno *et al.*, 2000). When using the tblastn program, the "nr" and "dbest" databases were searched. The "nr" nucleotide database, which is selected by default, includes all GenBank (Benson *et al.*, 2000), EMBL (Stoesser *et al.*, 2001), DDBJ (Tateno *et al.*, 2000), and PDB (Berman *et al.*, 2000) complete nucleotide sequences. Expressed sequence tags (ESTs) from a variety of organisms are found in the "dbest" (Boguski *et al.*, 1993) database. All other search parameters were left at the default settings and, as a result, a gapped alignment was produced, an expect (E) value of 10 was used, the filter for low complexity was activated, the BLOSUM62 matrix was selected, and 100 description and 50 alignments were returned.

Full-length or partial sequences corresponding to putative $P_{1B}$-ATPases were identified by browsing the descriptions and alignments that were found in the BLAST results. A sequence was selected for further analysis when it shared a high level of sequence similarity with the query sequences and contained at least one Type IB motif or when its description suggested it might be a sequence of interest. The protein and nucleotide records for the sequences of interest were initially retrieved using the links built into the BLAST results and were later accessed using the GI identifier.

The BLAST family of search programs will be used throughout this study. The search programs that will be used include: the blastp and tblastn programs that are described above, the blastn program that compares a nucleotide query against the nucleotide sequence database, the blastx program that compares a nucleotide query translated in all reading frames against the protein sequence database, the tblastx program that compares a nucleotide query translated in all reading frames against the nucleotide sequence database translated in all reading frames, and the Pattern-Hit Initiated Basic Local Alignment Search Tool (PHI-BLAST) that retrieves matches based on sequence similarity and the presence of a pattern selected by the user. Unless otherwise stated, the parameters that were used for all of the BLAST searches performed in this study were similar to those outlined above.

#### 2.2.1.3 Analysis of sequences that were identified in the literature and sequence database searches

##### 2.2.1.3.1 Identification of P-type and Type IB motifs

The protein sequences identified in the search for $P_{1B}$-ATPases were examined to identify the P-type (Section 1.1) and Type IB (Section 1.4) amino acid motifs that have been discussed previously. The P-type and Type IB motifs were first identified in the protein sequences using the DNAMAN multiple alignment tool. The CadA (AAB59154/150719), CopB (AAA61836/290643), Ccc2p (AAC37425/538515), ATP7A (AAA35580/179253), ATP7B (AAB52902/1947035), and identified plant amino acid sequences were aligned using the default settings of the optimal alignment option. When appropriate, the alignments were manually adjusted and the sequences were individually searched to identify any motifs that were missed during the alignments. P-type and Type IB motifs that are present in the ESTs were identified by examining the alignments that were returned in the tblastn searches.

#### 2.2.1.3.2 Predicting substrate specificity

The full-length amino acid sequences were analyzed to determine the group, $Cu^{2+}$- or $Cd^{2+}$-ATPases, with which the putative $P_{1B}$-ATPase sequences shared the greatest similarity. The homology tree feature, available from the output button of the DNAMAN multiple alignment results, was used to produce a crude phylogenetic tree. The blastp program was then used, without limiting the organism or taxonomic class, to identify the sequences available in the database that shared the highest similarity with the identified plant sequences.

#### 2.2.1.3.3 Grouping expressed sequence tags and full-length sequences

The blastn program was used to group the identified ESTs with the full-length sequences, if available, from which they were derived. These blastn searches were limited to the "nr" database and the plant species from which the EST had been derived. The blastn program was also used to identify additional ESTs that had been derived from the coding sequences of the full-length sequences that were identified in the database search. These blastn searches were limited to the "dbest" database and the appropriate plant species.

#### 2.2.1.3.4 Hydrophobicity

Hydrophobicity profiles were created and analyzed for the proteins that had been identified in the search for $P_{1B}$-ATPases from plants. The hydrophobicity profile option of the DNAMAN program was used with a window of 19 amino acids. The amino acid segments that likely represent membrane-spanning domains were identified according to the findings of Kyte and Doolittle (1982). These authors suggest that when the hydropathy of a given 19-residue segment averages greater than +1.6, there is a high probability that it will be one of the sequences in a membrane bound protein that spans the membrane. Hydrophobicity profiles for the *S. cerevisiae* Ccc2 (L36317/538514; $P_1$-

ATPase), *Schizosaccharomyces pombe* Pma2 (M60471/173430; $P_2$-ATPase), and *Escherichia coli* KdpB (K02670/2772547; $P_3$-ATPase) proteins were produced to illustrate the profiles that are characteristic of the $P_1$-, $P_2$-, and $P_3$-ATPases described by Lutsenko and Kaplan (1995).

## 2.2.2 Cloning of a *Brassica napus* cDNA that encodes a $P_{1B}$-ATPase

### 2.2.2.1 General techniques

Many of the general molecular biology techniques that were used throughout this study were as described by Sambrook *et al.* (1989), the manufacturers of various products, or the literature. Under these circumstances, the protocol will not be described in detail; instead, the reader will be referred to the original protocol and informed of any alterations that were made to the protocol. Other protocols used in the current study, which, were optimized and passed down through a number of different researchers and will be described in detail. The composition of common solutions and media mentioned in these protocols can be found in Sambrook *et al.* (1989).

#### 2.2.2.1.1 Isolation of nucleic acids from plant tissue

QIAGEN (Mississauga, Ontario, Canada) kits were used to isolate nucleic acids from plant tissue. Total DNA was isolated using the DNeasy™ Plant Mini Kit (Catalog number 69104) and the protocol found in the included handbook. Total DNA was isolated from five-day-old seedlings of either *Arabidopsis* ecotype C24 or *B. napus* cultivar Westar. The RNeasy® Plant Mini Kit (Catalog number 74904) and the Plant and Fungi Protocol in the RNeasy® Mini Handbook were used to isolate total RNA. The optional centrifugation step in a clean 2 ml collection tube was performed during the isolation of total RNA. Total RNA was isolated from *B. napus* cultivar Westar mature leaf tissue.

#### 2.2.2.1.2 Polymerase chain reaction primers

Polymerase chain reaction (PCR) primers were designed with the aid of the DNAMAN and GeneTool programs. The primers that were selected satisfied the General Guidelines for Standard PCR Primers outlined by QIAGEN in the *Taq* PCR Handbook that is included with the QIAGEN *Taq* DNA Polymerase Kit (Catalog number 201203). Primers were produced by one of three groups: the Molecular Biology Services Unit, Department of Biological Sciences, University of Alberta; the DNA Lab, Department of Biochemistry, University of Alberta; or GibcoBRL (Burlington, Ontario, Canada). Primers were suspended in 250 µl of water (Molecular Biology Services Unit and DNA Lab) or 10mM Tris, pH 8.0 (GibcoBRL) and stored in aliquots of 50 µl. A list of the primers used in the current study is available in Table 2.1.

#### 2.2.2.1.3 Polymerase chain reactions

Unless stated otherwise, the following conditions were constant for all PCRs and the solutions were from the QIAGEN *Taq* DNA Polymerase Kit (Catalog number 201203). The reactions were initially setup as master mixes that contained all of the PCR components (0.5 µM of each primer, 1x PCR Buffer, 0.4x Q-Solution, 1.6 mM dNTP mix, 2.5 mM $MgCl_2$, and *Taq* DNA polymerase) except the template. The stock solution of dNTP mix (40 mM) was made from the Amersham Pharmacia Biotech (Baie d'Urfé, Québec, Canada) Ultrapure dNTP Set (Catalog number 27-2035-01). All PCRs in which the products were to be used for further experimental work contained 0.625 units/reaction of *Taq* DNA Polymerase produced by QIAGEN. Any PCR reactions that were setup solely for diagnostic purposes used 0.25 µl/reaction of *Taq* DNA polymerase made in the lab of Dr. M. Pickard (University of Alberta, Edmonton, Alberta). The volume of each master mix was at least 10% greater than required for all of the samples and controls. The appropriate volume of master mix was distributed into 200 µl PCR tubes, and the appropriate template was added and mixed into the contents of each tube. The final volume of all PCRs was 25 µl.

The GeneAmp® PCR System 9600 or GeneAmp® PCR System 9700 (ramp speed set at 9600) was used for all PCRs. In general, the annealing temperature that was used in the PCR programs was 5° C below the average $T_m$ of the primers. The denaturation and extension temperatures used in all PCR programs were 94° C and 72° C, respectively. The first step of every standard PCR was a denaturation period that was 2 minutes long and the final step was an extension period of 5 minutes. All reactions were held at 4° C once complete. The initial denaturation period was followed by 35 denaturation/annealing/extension cycles in which each period was at least 1 minute long. When PCR products longer than 1000 bp were anticipated, the extension period within these cycles was increased by a minute, or portion thereof, for each additional 1000 bp.

Precautions were taken to limit the potential for contamination of the PCRs and to detect any contamination that did occur. All PCR master mixes were setup using a fresh tube of sterile Milli-Q water. Although a fresh tube of the stock solutions was not used for each master mix, the stock solutions were divided into smaller aliquots. The master mixes were setup in a DNA-free environment using aerosol resistant tips and pipettors dedicated only to setting up PCR master mixes. All microfuge tubes were also from a sterile stock dedicated to making PCR master mixes. A negative control that contained no template was included in every set of reactions. When appropriate, an additional negative control was included that contained a template, such as an empty vector, that was not expected to produce a product.

The reverse transcriptase-polymerase chain reaction (RT-PCR) is a specialized PCR that uses RNA as a template and produces cDNA products. The first step of the RT-PCRs that were used in this study used the Gibco SUPERSCRIPT™ II RNase H⁻ Reverse Transcriptase Kit (Catalog number 18064-014) to produce first strand cDNA. The protocol provided with the kit was followed using 5 µg of total RNA as the template and 500 ng of oligo dT primer or 2 pmoles of gene specific primer. The PCR step of the RT-PCR was setup with QIAGEN *Taq* DNA Polymerase or with QIAGEN HotStarTaq™ DNA Polymerase. The PCR reactions that included *Taq* DNA Polymerase were setup as

previously outlined, while the reactions that contained HotStarTaq™ DNA Polymerase required a 15 minute, 95° C heat activation step.

Colony PCR is a specialized PCR that was performed when the desired PCR template was in the cells within a colony instead of isolated nucleic acids. A sterile pippette tip was used to mix an isolated *E. coli* colony into 10 µl of sterile Milli-Q water that had been added to a 200 µl PCR tube. The cells that remained on the pippette tip were spread onto an appropriate agar plate that was later incubated. The PCR tube containing the water and *E. coli* cells was heated at 95° C for 5 minutes. Fifteen µl of master mix was then added to each PCR tube and the PCR was continued as previously outlined.

#### 2.2.2.1.4 Agarose gels

Agarose gels were made with 0.5x TBE and contained between 0.5 and 2.0% agarose. The percentage of agarose used in each gel varied with the expected size and desired resolution of the nucleic acid product(s); generally the concentration of agarose was increased as the expected size of the product(s) decreased. The agarose gels were all run in freshly diluted 0.5x TBE at a voltage that varied with the desired results. In general, a lower voltage was used when increased resolution or running time was desired. The maximum voltage applicable to the gel apparatus was used in most cases.

#### 2.2.2.1.5 Gel extraction

DNA fragments were purified from agarose gels using the QIAquick Gel Extraction Kit (QIAGEN Catalog number 28704). The QIAquick Gel Extraction Protocol Using a Microcentrifuge that is found in the QIAquick® Spin Handbook, which is provided with the kit, was followed. The optional Buffer QG wash was always preformed and the column was left to sit for 5 minutes after the addition of Buffer PE.

#### 2.2.2.1.6 Production of a T-vector for cloning polymerase chain reaction products

The adenine (A) residue that is added by *Taq* DNA polymerase to the 3' ends of PCR products was utilized in a TA cloning strategy. The T-vector, which contains a 3' thymine (T) overhang that facilitates cloning of the PCR products, was created using the pBluescript SK- vector. The pBluescript SK- vector was digested with *EcoRV* (4 µg pBluescript SK-, 36 units *EcoRV*, final volume of 20 µl, 37° C, 1 hour) to produce blunt ends to which a thymine (T) could be added. Next, dTTP (1 µl of 40 mM) and QIAGEN *Taq* DNA polymerase (1 µl) was added and the mixture was incubated at 70° C for 2 hours. The volume of the mixture was then adjusted to 100 µl with TE buffer (pH 8.0) and the DNA was extracted with phenol (100 µl), followed by chloroform:isoamyl alcohol (100 µl; 24:1). The DNA was then precipitated with 95% ethanol (500 µl), washed with 70% ethanol (500 µl), air dried, and resuspended in TE buffer (20 µl).

#### 2.2.2.1.7 Ligations

All ligations were setup using T4 DNA Ligase (Gibco BRL Catalog Number 15224-017). The ligation reactions were set up as outlined in the Rapid Ligation Protocol for Plasmid Cloning of DNA Fragments that is found in the product insert. Although a five-minute incubation is reported to be sufficient, a minimum incubation period of 1 hour was allotted. The appropriate controls were included whenever possible. For example, negative controls were setup in which only vector or insert DNA was added. A reaction that contained vector DNA that was capable of recircularizing was also setup as a positive control.

#### 2.2.2.1.8 The production of competent DH5∝ *Escherichia coli* cells

Competent cells were produced by a modified hexamminecobalt chloride/DMSO method. An LB plate was streaked from a DH5∝ glycerol stock and incubated overnight at 37° C. An isolated colony from this plate was used to inoculate 5 ml of LB liquid

medium that was then incubated overnight at 37° C and 250 RPM. One hundred µl of this overnight culture was used to inoculate 125 ml of SOB liquid medium in a 500 ml flask. The inoculated SOB was then grown at 25° C and 125 RPM until it reached an $OD_{600}$ of 0.4 (~10 hours). The culture was rapidly cooled on ice, placed in 2, 50 ml tubes, and centrifuged 10 minutes at 3000 x g and 4° C. The pellets were resuspended in a total of 80 ml of fresh FSB (see Sambrook *et al.*, 1989) and centrifuged for 10 minutes at 3000 x g and 4° C. The pellets were then resuspended in a total of 12.5 ml of FSB and pooled. The pooled mixture was swirled as 875 µl of DMSO was added. This mixture was divided into 400 µl aliquots that were frozen in a dry ice-ethanol bath and stored at -80° C.

A number of precautions were taken to help produce highly competent cells. All glassware was washed, rinsed thoroughly, and sterilized before use. All reagents and medium were prepared immediately before use. When appropriate, all solutions and tubes were chilled before use. Whenever possible manipulations were performed in the laminar flow hood.

#### 2.2.2.1.9 Transformation of competent DH5∝ *Escherichia coli* cells

Both ligation mixtures and intact plasmids were transformed into competent DH5∝ *E. coli* cells. Competent cells were thawed on ice and 4 µl of the ligation mixture or 1 ng of intact plasmid was added to the cells that were then left on ice for 15 minutes. The cells were exposed to a 42° C heat shock for 90 seconds and then transferred to ice for 2 minutes. One ml of LB broth was added to the cells and they were incubated with gentle shaking at 37° C for 45 minutes. Various volumes of the transformation mixture were spread onto LBA plates that were then incubated overnight at 37° C. When a pBluescript SK-based vector was being transformed into the *E. coli*, blue/white selection was possible and 40 µl of 20 mg/ml X-gal and 4 µl of 0.1 M IPTG was spread onto the plates before the transformation mix was added. Blue/white selection was not possible with the pYES3-based vector. A negative control containing only competent cells and a

positive control containing 1 ng of the appropriate empty vector were always setup. When applicable, the various ligation controls were also transformed into the competent cells.

#### 2.2.2.1.10 Preparation of plasmid DNA

Plasmid DNA that was to be used for diagnostic purposes was isolated using the alkaline lysis miniprep plasmid purification method described by Sambrook *et al.* (1989), while plasmid DNA that was to be used for further experimental work was prepared using the QIAprep Spin Miniprep Kit (QIAGEN Catalog number 27104). The QIAprep Spin Miniprep Kit Protocol Using a Microcentrifuge, which is found in the QIAprep miniprep handbook, was followed. Plasmid DNA was isolated from 5 ml overnight cultures and the optional Buffer PB wash was always preformed. In order to reduce the amount of salt in the eluate, columns were left to sit for 5 minutes after the addition of Buffer PE.

#### 2.2.2.1.11 Restriction enzyme digest reactions

All restriction enzymes were from Amersham Pharmacia Biotech. Unless otherwise stated, the restriction digests were setup using 10x One-Phor-All Buffer PLUS ($OPA^+$) diluted to 0.5x, 1x, or 2x. The amount of $OPA^+$ was determined using the $OPA^+$ activity guide provided by Amersham Pharmacia Biotech. Restriction digests were usually setup by making a master mix containing all components of the reaction except the DNA, allocating this mix out, and adding the DNA to the individual reactions. The final volume of the individual restriction reactions was usually 20 µl.

Double and triple digests were performed in the fewest steps possible. When all of the enzymes used the same strength of $OPA^+$, single reactions were setup with all of the enzymes. If one or more of the enzymes required a different strength of $OPA^+$ the first reaction was usually setup for the enzyme that required the least $OPA^+$. The first enzyme was then heat inactivated, the reaction conditions were adjusted, and the other

enzyme(s) were added. At times an enzymes could not be heat inactivated or it was undesirable for the first digestion to be preformed with the enzyme that requires the lowest concentration of $OPA^{+}$. When this was the case, the partially digested DNA was purified by gel extraction and the purified DNA was then digested with the remaining enzyme(s).

In 2001, Amersham Pharmacia Biotech introduced new restriction enzyme buffers (Buffers L, M, H, K, and T). These buffers were used for the restriction digests that were necessary to sub-clone the *Brassica napus* cDNA from the pBluescript-cDNA vector into the pYES3 vector. These restriction digests were setup based on the Activity Chart for Restriction Enzymes from Amersham Pharmacia Biotech. Whenever possible double digest were performed in a single buffer.

#### 2.2.2.1.12 Sequencing

All products were cloned prior to sequencing and the appropriate plasmid was purified using the QIAprep Spin Miniprep Kit. The Molecular Biology Services Unit, Department of Biological Sciences, University of Alberta or the DNA Lab, Department of Biochemistry, University of Alberta performed all sequencing reactions. In certain situations a product was not completely sequenced, these products were instead partially sequenced using only the T7 and/or T3 primers, which flank an insert that has been cloned into the multiple cloning site of the pBluescript SK- vector. These sequencing reactions usually provided at least 300 bp of clear insert sequence from each primer. To promote consistency when an insert is only partially sequenced, the sequencing results that are discussed will be limited to the first 300 bp of insert sequence provided by each primer. The sequences were analyzed by BLAST searches and/or a comparison to the appropriate region of a sequence of interest using the multiple alignment tool of the DNAMAN program.

#### 2.2.2.1.13 The identification of plasmids harboring an appropriate insert

Following a transformation it is necessary to examine the various transformants to identify those that contain a vector carrying the appropriate insert. In most cases this examination involved the purification and analysis of plasmid DNA. A preliminary analysis of the isolated plasmids was used to determine which of the isolated plasmids were likely to contain the insert of interest. This preliminary analysis varied with the size of the insert, and usually involved a size comparison of the isolated plasmids or a restriction analysis. Colony PCR, which allows a large number of transformants to be screened, was employed when an insufficient number of transformants harboring the vector of interest were identified by the preliminary analysis. Plasmid DNA was then isolated from the transformants that were identified by the colony PCR screen. Plasmids that were likely to contain the appropriate insert, as determined by the preliminary analysis or colony PCR, were subjected to further analysis that involved a restriction analysis, if not already done, a PCR analysis, and/or sequencing. The PCRs were setup with product specific primers and/or primers, such as T7 and T3, that flank the multiple cloning site.

#### 2.2.2.1.14 Glycerol stocks

Glycerol stocks were created to maintain each *E. coli* strain that had been transformed with a vector of interest. All glycerol stocks were made by mixing 2.25 ml of a sterile glycerol:LB liquid medium solution (1:1) with 5 ml of an overnight culture that had been initiated in the appropriate liquid medium, usually LBA, using an isolated colony. The mixture was allocated into sterile microcentifuge tubes, frozen with liquid nitrogen, and stored at –80° C.

#### 2.2.2.1.15 $^{32}$P-labeling of DNA fragments

The Amersham Pharmacia Biotech Oligolabelling Kit (Catalog number 27-9250-01) was used to label DNA fragments that had been purified by gel isolation. The

protocol included with the kit was followed. Amersham Pharmacia NICK™ Columns (Catalog number 17-0855-01) were used to separate the $^{32}$P-labeled DNA from the unincorporated $^{32}$P-labeled nucleotides. The protocol included with the NICK™ Columns was followed using TE buffer (pH 8.0).

### 2.2.2.2 Amplification of polymerase chain reaction products from plant putative $P_{1B}$-ATPases

The PAA1, RAN1, HMA1, HMA4, CadA (AAB59154/150719), CopB (AAA61836/290643), Ccc2p (AAC37425/538515), ATP7A (AAA35580/179253), and ATP7B (AAB52902/1947035) amino acid sequences were examined to identify conserved regions that could be used to design PCR primers suitable for the amplification of fragments from known or unknown $P_{1B}$-ATPases. *Arabidopsis* sequences from the databases were then used to predict the potential genomic and coding sequence products that might be produced using primers that were based on the identified regions. The *Arabidopsis PAA1*, *RAN1*, *HMA1*, and *HMA4* sequences were used to predict products that would correspond to putative $P_{1B}$-ATPases. Potential products from sequences that belonged to P-type ATPase subfamilies other than the Type IB subfamily were predicted using the full-length *Arabidopsis* sequences *ACA3*/U93845/1943750, *PEA1*/L08468/509809, *AHA2*/SwissProt P19456, *AHA3*/SwissProt P20431, *AHA9*/X73676/471289, *AHA10*/ S74033/765353 that had been identified by Axelsen and Palmgren (1998).

A set of partially degenerate primers was designed using the *Arabidopsis PAA1* and *RAN1* sequences. The sense, PCR-S, and antisense, PCR-A, primers were based on the amino acid sequences DKTGTL and GDGIND, respectively. The sequence of these primers can be found in Table 2.1 and the position of these primers, relative to the *PAA1* and *RAN1* coding sequences, is illustrated in Figure 2.1. Polymerase chain reactions (PCRs) were set up using the PCR-S/PCR-A primer pair and *Arabidopsis* or *B. napus* genomic DNA. Selected products were cloned and partially sequenced using the T7 and T3 primers.

### 2.2.2.3 Production of polymerase chain reaction-based probes that will be used to screen a *Brassica napus* cDNA library

A variety of PCR-based probes (Figures 2.2 and 2.3) were produced to screen an amplified *B. napus* cDNA library. The sequences of the primers that were used to produce these probes are reported in Table 2.1. Figures 2.2 and 2.3 depict the position of these primers relative to the *RAN1*, *PAA1*, and *HMA1* coding sequences.

The first probes, Probe1 and Probe2, were derived from fragments of the cloned *B. napus* PCR products PCR-Bn1 and PCR-Bn2. The PCR-Bn1 and PCR-Bn2 partial sequences were used to design the Probe1-S/Probe1-A and Probe2-S/Probe2-A primer pairs, respectively. Polymerase chain reactions (PCRs) were then setup using these primer pairs and the appropriate *B. napus* PCR product, PCR-Bn1 or PCR-Bn2. Polymerase chain reaction (PCR) products that were produced by these reactions were cloned and later liberated from the pBluescript SK- vector by *EcoRI / HindIII* restriction digests. The appropriate restriction fragments were then gel isolated and labeled to produce Probe1 (from PCR-Bn1 and Probe1-S/Probe1-A) and Probe2 (from PCR-Bn2 and Probe2-S/Probe2-A).

The next probes, Probe3 and Probe4, were based on regions of the *RAN1* and *PAA1* coding sequences that were close to the 5'-end of each coding sequence. The *RAN1* and *PAA1* coding and genomic sequences were used to design the Probe3-S/Probe3-A and Probe4-S/Probe4-A primer pairs, respectively. A blastn search against the *Arabidopsis* "dbest" database was performed to ensure that the predicted *Arabidopsis* PCR products did not contain any sequences that are common in the *Arabidopsis* "dbest" database. Polymerase chain reactions (PCRs) were then setup using these primer pairs and *Arabidopsis* or *B. napus* genomic DNA. Selected products were cloned and completely sequenced. The Probe3-Bn1 and Probe4-Bn1 products were liberated from the pBluescript SK- vector using *EcoRI / HindIII* and *PstI / HindIII* restriction digest reactions, respectively. The appropriate restriction fragments were then gel isolated and labeled to produce Probe3 and Probe4.

An updated version of the *HMA1* genomic entry (*HMA1*/Z99707/**4376087**), which was annotated with revised coding and protein sequences, was used to design another two pairs of PCR primers. The first pair of PCR primers, Probe5-S and Probe5-A, was based on a 5'-region of the *HMA1* sequence that was limited to the first exon and did not contain any regions of the *Arabidopsis* coding sequence that are common in the *Arabidopsis* "dbest" database. The second set of PCR primers, Probe6-S and Probe6-A, was based on sequences flanking the nucleotides that encode the SPC tripeptide, which had been identified in place of the CPx motif. The product that was predicted for these primers did not contain any sequences that are common in the *Arabidopsis* "dbest" database, but did contain an intron, as a desirable primer pair that did not span an intron could not be identified. Polymerase chain reactions (PCRs) were setup using the Probe5-S/ Probe5-A and Probe6-S/Probe6-A primer pairs and *Arabidopsis* or *B. napus* genomic DNA. Products from the resulting bands were cloned and sequenced. The Probe6-Bn2 product was liberated from the pBluescript SK- vector using the *HindIII* and *PstI* restriction enzymes. The appropriate restriction fragment was then gel isolated and labeled to produce Probe6.

#### 2.2.2.4 Screening the amplified *Brassica napus* cDNA library

Dr. Isobel Parkin at Agriculture and Agri-Food Canada (AAFC) kindly provided a *B. napus* cDNA library for use in this study. The library was generated from five-day-old, etiolated, *B. napus* seedlings from the AAFC double haploid line DH12075 and had been created using the Stratagene Uni-ZAP® XR Cloning Kit (Catalog number 200450). Excised pBluescript SK- vectors would, therefore, consist of a cDNA inserted into the *EcoRI* and *XhoI* sites. Upon arrival, the library was titered and amplified. The amplified library was then titered and used for screening. The titering, amplification, and plating stage of the screening procedure were performed according to the Instruction Manual supplied by Stratagene with the Uni-ZAP® XR Cloning Kit. Amersham Pharmacia Hybond-N Nylon Membranes Optimized for Nucleic Acid Transfer (Catalog numbers RPN137N and RPN82N) were used for the plaque lifts, which were performed as outlined in the Protocol Book that was supplied with these membranes. The DNA was

fixed to the membrane by baking for 2 hours at 80°C. The prehybridization, hybridization, and washing steps were also performed as outlined in the protocols supplied with the Hybond-N Nylon Membranes. Positives were cored and excised as described in the Stratagene Instruction Manual that is provided with the Cloning Kit. A *KpnI / XbaI* restriction digest was used to liberate the cDNA fragment(s) from the excised pBluescript SK- vector in order to determine their size. Select positives were further analyzed by partial sequencing using the T3 and T7 primers.

#### 2.2.2.5 Isolation of the 5'-coding sequence that is missing from the ~2000 bp library clone by reverse transcriptase-polymerase chain reaction

Three primers (Figure 2.4) were designed for use in a RT-PCR strategy that was intended to amplify the 5'-coding region that corresponds to the ~2000 bp cDNA that had been isolated from the amplified *B. napus* cDNA library. The 5'-untranslated region from the *RAN1* mRNA/CDS record (AF091112/4760379) was used to design the first primer, which was named RT1-S. The second primer, RT2-S, was based on the 5'-untranslated region and several of the initial nucleotides found in the *RAN1* mRNA/CDS record (AF091112/4760379). An antisense primer, RT1/2-A, was designed that was compatible with both the RT1-S and RT2-S primers. This antisense primer was based on the sequencing information that had been obtained from a preliminary sequencing reaction of the ~2000 bp cDNA. Two other primers, Probe1-A and RT3-A, were used in the RT-PCR (Figure 2.4). The Probe1-A primer had been designed previously to produce Probe1, while the RT3-A primer had been designed for a PCR that is not discussed. The Probe1-A, RT1/2-A, and RT3-A primers were use for first strand cDNA synthesis. This first strand cDNA was then used in PCRs that paired either the RT1-S or the RT2-S primer with the RT1/2-A primer. Fragments from the RT-PCR bands were cloned and analyzed using the restriction enzymes *PstI*, *RsaI*, *SacI*, and *XhoI*. Select RT-PCR products were then partially sequenced using the T7 and T3 primers. The sequences of all of the primers that were used in the RT-PCR are listed in Table 2.1.

### 2.2.2.6 Production of a full-length *Brassica napus* cDNA

A series of restriction digest reactions and gel isolations were used to prepare the pBluescript SK- vector (Figure 2.5A), RT-Bn2 or RT-Bn3 product (Figure 2.5C), and ~2000 bp cDNA (Figure 2.5B) for the ligation reaction. In addition to the *SacI* site that was to be used in the ligation strategy, the RT-Bn2 product, RT-Bn3 product, and ~2000 bp cDNA each contained a second *SacI* site in the pBluescript SK- multiple cloning site. These additional *SacI* sites could potentially complicate the restriction digests and ligation reaction. A series of restriction digests and gel isolations were, therefore, used to prepare the RT-Bn2 product, RT-Bn3 product, and ~2000 bp cDNA for the ligation reaction. The pBluescript SK- vector was digested with *XhoI* and *XbaI* and the appropriate fragment was gel isolated. The ~2000 bp cDNA and RT-Bn2/RT-Bn3 products were liberated from the pBluescript SK- vector by *ApaI / XbaI* and *XhoI / XbaI* digests, respectively. The appropriate fragments from these restriction digest reactions were then gel isolated and digested with *SacI*. The fragments that were to be used in the ligation reaction were then gel isolated. The ligation reaction, therefore, contained the pBluescript SK- vector with *XhoI / XbaI* ends, the RT-Bn2 or RT-Bn3 RT-PCR fragment with *XhoI / SacI* ends, and the ~2000 bp cDNA with *SacI / XbaI* ends.

### 2.2.2.7 Sequencing and analysis of the full-length *Brassica napus* cDNA

A stock of the pBluescript SK- vector containing the full-length *B. napus* cDNA (pBluescript-cDNA; Figure 2.6) was prepared using the QIAprep Spin Miniprep Kit (Section 2.2.2.1.10). The concentration of the pBluescript-cDNA plasmid that was yielded directly by the QIAprep Spin Miniprep Kit was too low for sequencing and a SpeedVac (Savant SC100) was used to concentrate the stock. To avoid excess salt in the concentrated stock, the DNA was eluted from the QIAquick membrane with an elution buffer that had been diluted 1/20 with sterile Milli-Q water and the plasmid stock was never concentrated more than 10x. Aliquots of the concentrated stock were sent for

sequencing using the T7, Seq1-S, Seq2-S, Seq3-A, Seq4-S, Seq5-S, Seq6-S, Seq7-A, Probe1-S, Seq8-A, Seq9-A, and T3 primers. With the exception of the T7 and T3 primers, these primers were based on sequencing results from the full-length *B. napus* cDNA or the PCR-Bn1 product. The sequences for these primers are listed in Table 2.1. The position of these primers, relative to the full-length *B. napus* cDNA, is illustrated in Figure 2.7. The chromatographs that corresponded to the sequencing results were scrutinized to ensure that all of the sequencing peaks were clear. If a peak was unclear the appropriate region was resequenced. The translation tool in the DNAMAN program was used to predict the amino acid sequence from the *B. napus* cDNA sequence.

## 2.3 Results

### 2.3.1 Identification and analysis of plant sequences that correspond to putative $P_{1B}$-ATPases

#### 2.3.1.1 Literature search

A single journal article, which described an *Arabidopsis* cDNA that was believed to encode a $P_{1B}$-ATPase (Tabata *et al.*, 1997), was identified by the literature search. This putative $P_{1B}$-ATPase was named P-type ATPase of *Arabidopsis* (PAA1). The amino acid sequence encoded by the cDNA contained all of the P-type (Section 1.1) and Type IB (Section 1.4) motifs that were discussed previously. The variable N-terminal metal binding region was represented by a single copy of the CxxC motif. An alignment of PAA1 with the protein database revealed that PAA1 shared the greatest sequence similarity with proteins that have been proposed to be $Cu^{2+}$-ATPases (Tabata *et al.*, 1997). No functional evidence was provided to support the proposed role of PAA1 as a functional transporter or to support a possible substrate.

Tabata *et al.* (1997) cloned the *PAA1* cDNA from a λgt11 cDNA library made from greening rosettes of *Arabidopsis thaliana* ecotype Columbia. The probe used to

screen this library was produced by PCR using degenerate primers and an *Arabidopsis* cDNA plasmid-bank as the template. The primers were based on the standard codons for the amino acid sequences DKTGTL and DG[TIV]ND[AS]. These amino acid sequence are found in many P-type ATPases and the primers were, therefore, not specific for $P_{1B}$-ATPases. Tabata *et al.* (1997) identified a group of PCR products that were 600-700 bp in length and another set of fragments that were larger than 1000 bp. The products that were larger than 1000 bp were assumed to be from P-type ATPases other than the $P_{1B}$-ATPases and were, consequently, not cloned. The 600-700 bp fragments were cloned. One of the cloned PCR fragments was found to encode an amino acid sequence that was similar to the corresponding region of P-type ATPases. The amino acid sequence that was predicted from this PCR fragment was more similar to the putative $Cu^{2+}$-ATPases PacS (43%) and ATP7B (40%) than to a human classical $Ca^{2+}$-ATPase (28%; Tabata *et al.*, 1997).

The cloning of the *PAA1* cDNA (Tabata *et al.*, 1997) confirms that a least one full-length putative $P_{1B}$-ATPase is encoded by the *Arabidopsis* genome. None of the ESTs that were identified by Axelsen and Palmgren (1998) match the *PAA1* coding sequence (Section 2.3.1.3.2) and it is, therefore, likely that the *Arabidopsis* genome encodes other $P_{1B}$-ATPase(s). The method that Tabata *et al.* (1997) used to clone the *PAA1* cDNA suggests that the $P_{1B}$-ATPases contain fewer nucleotides between the regions that encode the DKTGT[LIVM][TIS] and GDGxNDxP amino acid motifs and that this feature can be used in a cloning strategy designed to isolate $P_{1B}$-ATPases. In this strategy the DKTGT[LIVM][TIS] and GDGxNDxP amino acid motifs are used to design PCR primers and the products that have likely originated from $P_{1B}$-ATPases are selected based on their smaller size.

#### 2.3.1.2 Sequence database search: Basic Local Alignment Search Tool at the National Center for Biotechnology Information

Both nucleotide and protein records were identified in the database searches. The nucleotide entries were EST, mRNA/CDS, or genomic records. The EST and

mRNA/CDS entries were derived from single genes, while the genomic entries contained many genes. When discussing a single gene within a genomic sequence it is, therefore, necessary to locate the sequence of interest by, for example, searching for the genomic locus, if available, or the corresponding protein accession number or GI identifier. The EST entries contained only a nucleotide sequence, while the mRNA/CDS and genomic entries were annotated with single and multiple predicted amino acid sequences, respectively. The predicted amino acid sequences that were found within the mRNA/CDS or genomic entries were also usually found as separate protein records. These protein records, although they may have originated from a genomic sequence that encodes many proteins, contain only a single amino acid sequence.

Sequences in the NCBI databases are assigned an accession number and a GI identifier. The accession number is relatively constant and does not usually change when the information in a record is altered; as a result, the accession number provides a method of retrieving the most recent record for a sequence. A new GI identifier is, theoretically, assigned to a record every time a change is made and the GI identifier should, therefore, provide a method of retrieving the exact record being discussed, irrespective of revisions that have since been made. The sequences that were identified in this study will initially be described using the format: Gene/protein name (if available)/accession number/GI identifier. For simplicity this description will be shortened to the gene/protein name, if available, in later discussions. The full sequence description (Gene/protein name (if available)/accession number/GI identifier) will be reported anytime an updated version of the original entry is being discussed. Most of the putative $P_{1B}$-ATPases that were identified in the database search have now been assigned a name (Axelsen and Palmgren, 2001) and these names will be used throughout the study to describe the corresponding nucleotide (italicized) or protein sequence. It is, however, important to note that contemporary knowledge that is associated with these names (*PAA1*/PAA1, *RAN1*/RAN1, *HMA1*/HMA1, and *HMA4*/HMA4) may not have been available at the time of the searches.

A total of five plant sequences that encode putative $P_{1B}$-ATPases were identified in the database searches. Four sequence sets were identified from *Arabidopsis* and one sequence set was identified from *Glycine max* (soybean). The first *Arabidopsis* sequence records were the mRNA/CDS (*PAA1*/D89981/2668491) and protein (PAA1/BAA23769/2668492) entries supplied in conjunction with the PAA1 reference (Tabata *et al.*, 1997). The *Arabidopsis* Genome Initiative had not yet determined the genomic sequence for *PAA1*. The next two *Arabidopsis* sequence sets contained both genomic (*RAN1*/AC002342/2660661 genomic locus T19K24.18 and *HMA1*/Z99707/2464848) and protein (RAN1/AAC79141/2660670 and HMA1/2464854/2464854) records. The final *Arabidopsis* sequence contained only a genomic record (*HMA4*/AC002392/3176701 genomic locus T20K24.12) that was annotated with protein sequences. The *G. max* sequence set contained both a mRNA/CDS record (AF019115/2565258) and a protein (AAB81947/2565259) record.

Four plant ESTs that encode amino acids that shared a high level of homology with $P_{1B}$-ATPases and contained at least one Type IB motif (Section 1.4) were also identified. Single ESTs were identified from *Gossypium hirsutum* (cotton; AI054702/3325816) and *Oryza sativa* (rice; D42886/3107148), while two ESTs were identified from *Arabidopsis* (T43560/2758407 and Z33731/493320). Axelsen and Palmgren (1998) had also identified the first of these *Arabidopsis* ESTs (T43560/2758407).

The BLAST searches will be discussed in great detail as they offered an efficient and comprehensive means of identifying sequences corresponding to putative $P_{1B}$-ATPases from plants; it should, however, be noted that some of these sequences were also identified through a number of other searches (data not shown). For example, the Entrez database retrieval system (Schuler *et al.*, 1996), which is supported by NCBI, can be used to search the DNA and protein sequence databases using a text based strategy. The "related sequences" links associated with putative $P_{1B}$-ATPases provides an additional source of sequence descriptions that can be searched to identify other putative $P_{1B}$-ATPases. Finally, an *Arabidopsis* database can be accessed through The Institute for

Genomic Research homepage (TIGR; http://www.tigr.org) and searched using an Expanded Role Table that is based on the classification scheme used at the Munich Information Center for Protein Sequences (MIPS; Mewes *et al.*, 2000). Although these searches did reveal some of the putative $P_{1B}$-ATPases that had been identified by BLAST searches (data not shown), they were limited as they relied heavily upon the descriptions that have been assigned to a particular sequence.

#### 2.3.1.3 Analysis of sequences that were identified in the literature and sequence database searches

##### 2.3.1.3.1 Identification of P-type and Type IB motifs and the prediction of substrate specificity

An analysis of the *G. max* amino acid sequence (AAB81947/2565259) revealed that it is likely an incomplete $P_{1B}$-ATPase sequence. At 329 amino acids, the *G. max* protein is shorter than other $P_{1B}$-ATPases such as the 727 amino acid CadA (AAB59154/150719) or the 949 amino acid PAA1 (BAA23769/2668492). Furthermore, the DKTGT[LIVM][TIS] motif, which is required to classify a protein as a P-type ATPase, was not present and the TGE motif that is also found in the amino acid sequence of P-type ATPases was absent. The Type IB N-terminal metal-binding region and intramembranous CPx motif were also lacking. A sequence (GEGINDAP) similar to the P-type motif GDGxNDxP was present, as was the Type IB HP locus. The *G. max* sequence, therefore, contains only some of the amino acid sequences that have been suggested to be characteristic of P-type ATPases (Section 1.1) and/or members of the Type IB subfamily (Section 1.4). A blastp search, which was not limited to a particular organism or taxonomic class, revealed that HMA1 from *Arabidopsis* shared a high level of similarity with the *G. max* predicted protein. An amino acid alignment, performed using the DNAMAN program and the appropriate fragment from HMA1, revealed that these proteins are ~57% identical in the overlapping region. This relatively high sequence identity and the presence of a P-type and a Type IB motif suggests that this *G. max* sequence could represent a truncated HMA1 homologue.

P-type (Section 1.1) and Type IB (Section 1.4) motifs were identified in the *Arabidopsis* amino acid sequences. The PAA1 and RAN1 sequences contained all of the P-type and Type IB motifs. In these sequences, the metal binding region contained one (PAA1) or two (RAN1) copies of the CxxC motif. All but one of the P-type motifs and two of the three Type IB motifs could be identified in HMA1. The HMA1 amino acid sequence contained the GEGINDP sequence in place of the P-type motif GDGxNDxP. The amino acids GEGINDP had also been found in place of the GDGxNDxP motif in the *G. max* (AAB81947/2565259) partial sequence. A Pattern-Hit Initiated BLAST (PHI-BLAST) search revealed that this GEGxNDxP amino acid sequence does not appear to be present in any other putative P-type ATPases. The Type IB motifs that were found in HMA1 included the HP locus and a histidine-rich region in the N-terminus that could be considered a metal-binding region. The HMA1 amino acid sequence contained a SPC tripeptide in place of the remaining Type IB motif, CPx. A PHI-BLAST search revealed that putative $P_{1B}$-ATPases from a number of bacterial species (data not shown) also contained the SPC tripeptide in place of CPx motif. All of the P-type motifs and two of the three Type IB motifs could be identified in the HMA4 amino acid sequence. The HMA4 sequence contained the HP locus and the CPx motif; an N-terminal putative metal-binding domain could not, however, be identified. The HMA4 C-terminus does, nonetheless, contain twelve cysteine (C) pairs and a histidine-rich (H-rich) region that could potentially act as putative metal-binding domains (personal observations; Williams *et al.*, 2000).

The putative $P_{1B}$-ATPases identified in the database search were examined in an attempt to predict the putative substrate of each protein. A crude phylogenetic tree that was produced using the DNAMAN program (Figure 2.8) suggested that the HMA4 sequence clusters with CadA ($Cd^{2+}$-ATPase), while PAA1 and RAN1 cluster with the ATP7A, ATP7B, Ccc2p, and CopB ($Cu^{2+}$-ATPases). The HMA1 protein appears to be separate from all of the other proteins. A series of blastp searches that used the PAA1, RAN1, HMA1, and HMA4 amino acid sequences as queries revealed that the top matches for a particular query were all from the same group within the Type IB subfamily, either $Cd^{2+}$- or $Cu^{2+}$-ATPases. The PAA1 and RAN1 amino acid sequences

shared the greatest similarity with the $Cu^{2+}$-ATPases, while the HMA1 and HMA4 amino acid sequences were most similar to the $Cd^{2+}$-ATPases.

All of the *Arabidopsis* sequences contained the DKTGT[LIVM][TIS] motif as well as the other motifs that are found in most P-type ATPases (Section 1.1). Consequently, there is strong support suggesting that these *Arabidopsis* proteins belong to the P-type ATPase superfamily. Each *Arabidopsis* sequence also contained at least two of the Type IB motifs (Section 1.4) suggesting that the sequences might represent P-type ATPases from the Type IB branch of the superfamily ($P_{1B}$-ATPases). The phylogenetic tree and the blastp results suggest that PAA1 and RAN1 are likely $Cu^{2+}$-ATPases, while HMA1 and HMA4 might be members of the $Cd^{2+}$-ATPase division. The $Cd^{2+}$-ATPases reported in the literature have all been from prokaryotic organisms and the HMA1 and HMA4 might, therefore, represent the first eukaryotic $Cd^{2+}$-ATPases. It should, however, be noted that HMA1 and HMA4 sequences did not contain all of the typical Type IB motifs and could potentially be involved in the transport of a substrate that has not yet been reported for $P_{1B}$-ATPases.

Expressed sequence tags (ESTs) were identified that contained at least one Type IB motif and shared a high level of sequence similarity with the $P_{1B}$-ATPases. The *G. hirsutum* EST (AI054702/3325816) encoded the CPx and DKTGT[LIVM][TIS] motifs, while the *O. sativa* EST (D42886/3107148) encoded the HP locus. The first *Arabidopsis* EST (T43560/2758407) encoded the DKTGT[LIVM][TIS] motif and the HP locus. The second *Arabidopsis* EST (Z33731/493320) encoded a histidine-rich region.

#### 2.3.1.3.2 Grouping expressed sequence tags and full-length sequences

Full-length sequences were not identified when the *G. hirsutum* and *O. sativa* ESTs were used as blastn queries. The blastn searches did reveal that the first *Arabidopsis* EST (T43560/2758407) represented a portion of the *RAN1* nucleotide sequence and the second *Arabidopsis* EST (Z33731/493320) corresponded to a portion of the *HMA1* nucleotide sequence.

Only one (T43560/2758407) of the six ESTs that were reported by Axelsen and Palmgren (1998) was identified in the current study, which searched for ESTs that shared a high level of sequence similarity with $P_{1B}$-ATPases and contained a Type IB motif. The other ESTs that had been identified by Axelsen and Palmgren (1998) were used as queries in blastn searches. The N96835/1268954, T46552/2763235, and W43689/2748959 ESTs were found to correspond to the *RAN1* coding sequence. The Z33730/493319 EST was found to represent a portion of the *HMA1* coding sequence. The final EST, T403098/2597666, was not from a $P_{1B}$-ATPase, it was from a putative copper/zinc superoxide dismutase copper chaperone (AF061517/3108346). The amino acid sequence of some metallochaperones is similar to the N-terminus of some $P_{1B}$-ATPases (Huffman and O'Halloran, 2001).

The identification of ESTs from *G. hirsutum* and *O. sativa* provides support for the existence of $P_{1B}$-ATPases in plant species in addition to *Arabidopsis* and *G. max*. The fact that each of these ESTs shares a high level of sequence similarity with the $P_{1B}$-ATPases and contains a putative Type IB motif strengthens this argument. The presence of the P-type ATPase DKTGT[LIVM][TIS] motif in the *G. hirsutum* EST (AI054702/3325816) further reinforces the suggestion that it is from a P-type ATPase. In addition, blastx searches using either the *G. hirsutum* (AI054702/3325816) or the *O. sativa* (D42886/3107148) EST sequences as a query confirm that these ESTs share the greatest sequence similarity with the $P_{1B}$-ATPases. In contrast, a blastx search using the T403098/2597666 EST, which had been identified by Axelsen and Palmgren (1998) and was determined in this study to have originated from a sequence encoding a metallochaperone, suggested that this EST is more similar to the metallochaperones than to $P_{1B}$-ATPases. The two *Arabidopsis* ESTs (T43560/2758407 and Z33731/493320) that were found to contain Type IB motifs and shared sequence similarities with the $P_{1B}$-ATPases were both from complete *Arabidopsis* nucleotide sequences that were identified in the database searches. As a result, the tblastn "dbest" search did not provide evidence for additional putative $P_{1B}$-ATPases in *Arabidopsis*.

The *Arabidopsis* "dbest" database was searched, using the blastn program and the *PAA1*, *RAN1*, *HMA1*, and *HMA4* coding sequences as queries, to identify additional ESTs that corresponded to these sequences. The ESTs that were identified included those discussed above, as well as additional ESTs that had not been identified by Axelsen and Palmgren (1998) or the previous searches. There were no ESTs identified that corresponded to the *PAA1* mRNA/CDS entry. The *PAA1* coding sequence was, however, not in question as it was not predicted, but was instead determined from a cloned cDNA (Tabata *et al.*, 1997). There were no additional ESTs identified for the *HMA1* coding sequence. Additional ESTs were identified that had been derived from the *RAN1* (AA067528/1565877) and *HMA4* (AA728413/2747370, T04515/315675, H36471/905970, H36072/905571, H36075/905574, R90705/958245) coding sequences. Figures 2.9 and 2.10 provide a summary of all of the ESTs that were identified for the *RAN1*, *HMA1*, and *HMA4* coding sequences. These ESTs suggest that the *RAN1*, *HMA1*, and *HMA4* genes are being transcribed and confirm regions of the predicted coding sequences.

#### 2.3.1.3.3 Hydrophobicity

The determination of the exact number of transmembrane domains in P-type ATPases has been a matter of debate (Lutsenko and Kaplan, 1995) and no attempt was made to identify all of the transmembrane domains that likely occur in PAA1, RAN1, HMA1, and HMA4. The regions that had a high probability of representing a membrane-spanning domain (Kyte and Doolittle, 1982) were, however, highlighted in black (Figure 2.11). The overall hydrophobicity profiles were then compared to the three divisions outlined by Lutsenko and Kaplan (1995; Figure 2.12). The hydrophobicity profiles for the *S. cerevisiae* Ccc2 (L36317/538514; $P_1$-ATPase; Figure 2.12A), *Schizosaccharomyces pombe* Pma2 (M60471/173430; $P_2$-ATPase; Figure 2.12B), and *Escherichia coli* KdpB (K02670/2772547; $P_3$-ATPase; Figure 2.12C) proteins were produced to illustrate the profiles that are characteristic of these divisions ($P_1$-, $P_2$-, and $P_3$-ATPases; Lutsenko and Kaplan, 1995). The transmembrane domains that were identified by Lutsenko and Kaplan (1995), but did not have a high probability of

representing membrane spanning domains according to the findings of Kyte and Doolittle (1982), were highlighted in grey.

The N-terminal regions (prior to the TGE motif) of the PAA1, RAN1, HMA1, and HMA4 amino acid sequences contain a relatively high number of hydrophobic regions (Figure 2.11) and are, therefore, most similar to the $P_1$-ATPases (Figure 2.12A; note that the when the more stringent criteria of Kyte and Doolittle (1982) are used, it is not always possible to identify the four N-terminal transmembrane domains that Lutsenko and Kaplan (1995) proposed for $P_1$-ATPases). The PAA1, RAN1, and HMA1 sequences all exhibit a short C-terminus (following the GDGxNDxP motif) that contains a relatively low number of hydrophobic regions (Figures 2.11A, 2.11B, and 2.11C) and this is also characteristic of the $P_1$-ATPases (Figure 2.12A). In contrast, the C-termini (following the GDGxNDxP motif) of $P_2$- and $P_3$-ATPases (Figures 2.12B and 2.12C) is extended and contains a greater number of hydrophobic regions. Although the C-terminus (following the GDGxNDxP motif) of the HMA4 sequence is extended (Figure 2.11D), it is similar to the $P_1$-ATPases (Figure 2.12A) in that it contains a relatively low number of hydrophobic regions. The hydrophobicity profiles of the PAA1, RAN1, and HMA1 sequences are most similar to those of the $P_1$-ATPases. The HMA4 sequence is also most similar to the $P_1$-ATPases; however, as was also observed at the amino acid sequence level, the HMA4 protein contains a C-terminus that is unique from that of other $P_1$-ATPases/$P_{1B}$-ATPases. It should be noted that the $P_1$-ATPase group identified by Lutsenko and Kaplan (1995) is equivalent to the $P_{1B}$-ATPase subfamily identified by Axelsen and Palmgren (1998). The hydrophobicity profiles, therefore, provide addition evidence suggesting that the PAA1, HMA1, HMA4, and RAN1 proteins are likely $P_{1B}$-ATPases.

## 2.3.2 Cloning of a *Brassica napus* cDNA that encodes a $P_{1B}$-ATPase

### 2.3.2.1 The amplification of polymerase chain reaction products from plant putative $P_{1B}$-ATPases

The amino acid sequences of a group of $P_{1B}$-ATPases were examined to identify regions that could be used to produce PCR primers suitable for the amplification of fragments from known or unknown $P_{1B}$-ATPases. Ideally these primers would be specific for $P_{1B}$-ATPases and would not amplify products from P-type ATPases that belong to other divisions of the superfamily. Since a PCR primer should be at least 18 nucleotides long and would, thus, require an amino acid motif of at least 6 residues, all of the Type IB motifs (Section 1.4) were too short to be of any use in designing PCR primers. In addition, the database searches revealed that, of the Type IB motifs, only the HP locus was found in all of the putative $P_{1B}$-ATPases. The P-type ATPase motifs DKTGT[LIVM][TIS] and GDGxNDxP were longer and could be used to design PCR primers; unfortunately, these motifs are found in all P-type ATPases and primers that are based on these sequences could potentially produce fragments from all of the divisions of the P-type ATPase superfamily. As discussed previously (Section 2.3.1.1), it might, however, be possible to select PCR fragments originating from these primers and $P_{1B}$-ATPase templates by their smaller size.

*Arabidopsis* genomic and coding sequences that are found in the databases were used to predict potential products that could be produced using primers based on the DKTGT[LIVM][TIS] and GDGxNDxP motifs. The fragments that were predicted from the $P_{1B}$-ATPase coding sequences were smaller than those predicted for the coding sequence of other P-type ATPases (Table 2.2). This size difference was not observed when the corresponding genomic sequences were used to predict the products (Table 2.2). The number and size of introns in the predicted genomic products varied (Table 2.2) and, consequently, obscured the size difference that is observed in the coding

sequence products from $P_{1B}$-ATPases and other P-type ATPases. An examination of the size of products that are amplified using DKTGT[LIVM][TIS]- and GDGxNDxP-based primers could, therefore, provide a primary mechanism of identifying those fragments that are likely to have originated from $P_{1B}$-ATPases when this region is amplified from a coding sequence, but not when the PCR products are derived from the corresponding genomic sequence.

The predicted products were further analyzed to determine if the sequence of each fragment provided sufficient information to accurately predict the division of the P-type ATPase superfamily to which the product might belong. The HP locus was encoded by all of the fragments that originate from $P_{1B}$-ATPases, while the fragments from other divisions of the superfamily did not encode the HP locus. The HP locus could, therefore, be a reliable marker, when combined with other information, for $P_{1B}$-ATPases. When the predicted products were used as blastn queries, the highest matches were the original full-length sequences used to predict the sequences of the products. The blastn search tool could, therefore, be used to determine the full-length sequence, if available, that corresponds to a product. The full-length sequence could then be used to more accurately determine the division of the P-type ATPase superfamily to which the product and full-length sequence are likely to belong. When the predicted products were used as blastx queries, the top matches were all from the same division of the P-type ATPase family and this division was consistent with the putative classification of the full-length sequence that had been used to predict the PCR fragment. As a result, if a full-length sequence record is not available for a fragment, a blastx search could be used to assign the product to a division of the P-type ATPase superfamily. These results suggest that fragments produced with primers based on the DKTGT[LIVM][TIS] and GDGxNDxP motifs, when sequenced and appropriately analyzed, should be sufficient to predict the division of the P-type ATPase superfamily to which the corresponding complete sequence is likely to belong.

The DKTGT[LIVM][TIS] and GDGxNDxP motifs were used to design a set of partially degenerate PCR primers. The decision to limit the degeneracy was based on the

realization that a size selection strategy could not be used to identify genomic products that were likely amplified from $P_{1B}$-ATPases. The *Arabidopsis PAA1* and *RAN1* sequences were selected for use in designing the PCR primers. The PAA1 and RAN1 proteins contain all of the P-type and Type IB motifs and display hydrophobicity profiles similar to other $P_1$-ATPases/$P_{1B}$-ATPases. A substrate specificity can also be predicted for the RAN1 and PAA1 proteins with a relatively high level of confidence. Although the resulting primers are likely to be biased towards *PAA1*, *RAN1*, and any homologues of these genes, it is still possible that products from novel $P_{1B}$-ATPases might be identified. Since conserved amino acids were used to design the primers, the probability that the primers would work in plant species other than *Arabidopsis* was increased.

Polymerase chain reactions (PCRs) were setup using the partially degenerate primers, which were named PCR-S and PCR-A, and *Arabidopsis* genomic DNA as a template. Two clear bands of ~1050 and ~1400 bp (Figure 2.13) were produced by the PCR. The ~1050 bp PCR fragment was approximately the same size as had been predicted for the *RAN1* genomic PCR product (1038 bp; Table 2.2). Cloning, partial sequencing, and further analysis of this ~1050 bp product, which was named PCR-At1, suggested that it was amplified from the *RAN1* genomic sequence (Table 2.3). The *PAA1* genomic sequence was not available in the databases and the size of a genomic PCR product could not be predicted. Since the *PAA1* coding sequence had been used to limit the degeneracy of the primers, it was probable that the ~1400 bp fragment had been amplified from the *PAA1* genomic sequence. Cloning, partial sequencing, and analysis of the ~1400 bp fragment, which was named PCR-At2, suggested that this PCR product was likely amplified from the *PAA1* genomic sequence (Table 2.3).

The above results confirm that the PCR-S/PCR-A primer pair is functional and suggested that when an *Arabidopsis* template is used these primers are biased towards the *PAA1* and *RAN1* sequences from which they were designed. Since the PCR-S and PCR-A primers were based on highly conserved amino acids, it was likely that these amino acids and the nucleotides they are encoded by might be conserved in other plants. All of the full-length sequences that were available in the literature or sequence database for

putative plant $P_{1B}$-ATPases were from *Arabidopsis*. The PCR-S/PCR-A primers were, consequently, used in combination with templates from plant species other than *Arabidopsis*. *Triticum aestivum* and *B. napus* were selected because they are both important crop species and resources were available to work with these species. In addition, *B. napus* and *Arabidopsis* are closely related and the probability of success using the PCR-S and PCR-A primers with *B. napus* nucleic acids as a template was, therefore, higher than with templates from other more distantly related plant species. Although a number of PCRs and RT-PCRs were set up with the PCR-S/PCR-A primer pair and *T. aestivum* templates (data not shown), distinct products were not produced. Reverse transcriptase-polymerase chain reactions (RT-PCRs) that were setup with *B. napus* templates (data not shown) were also unsuccessful.

Polymerase chain reactions (PCRs) were setup using the partially degenerate primers PCR-S and PCR-A and *B. napus* genomic DNA. The PCR conditions, with the exception of the template, were identical to those that had been used for the *Arabidopsis* PCR. Under these conditions numerous bands were amplified from the *B. napus* genomic template (Figure 2.13). Since genomic DNA had been used as the template the smaller PCR products could not be assumed to represent fragments from $P_{1B}$-ATPases. Fragments from the most prominent band of the *B. napus* PCR (Figure 2.13; ~900 bp) were, therefore, cloned and it was subsequently discovered that this band actually contained at least two different PCR fragments, which were ~1000 and ~1250 bp in size. An analysis of the partial sequences that were determined for these cloned PCR fragments revealed that the ~1000 bp product, which was named PCR-Bn1, shared a high level of sequence similarity with *RAN1* (86 and 89% identity when coding sequences were compared), while the sequence of the ~1250 bp fragment, which was named PCR-Bn2, was similar to *PAA1* (87 and 86% identity when coding sequences were compared; Table 2.3). These cloned *B. napus* PCR products were, therefore, likely amplified from *RAN1* and *PAA1* homologues and were, consequently likely to be from $Cu^{2+}$-ATPases. Blastx searches in which the partial sequences were used as queries also suggested that the PCR-Bn1 and PCR-Bn2 products were from $Cu^{2+}$-ATPases. The HP locus was encoded by the partial sequences that were determined for the PCR-Bn1 and PCR-Bn2

products, further supporting the suggestion that these PCR products were amplified from $P_{1B}$-ATPases.

A high level of identity was observed when the PCR-Bn1 and *RAN1* predicted coding sequences were compared (86 and 89%; Table 2.3). This high level of sequence identity was observed across all regions of the coding sequence (Figures 2.14 and 2.15) and was not limited to regions of the coding sequence that encode amino acids that are conserved in all P-type ATPases and/or $P_{1B}$-ATPases. It is, therefore, likely that primers that are designed from any region of the *Arabidopsis* coding sequence will be successful when used in combination with a *B. napus* template. The level of identity that was observed when the PCR-Bn1 and *RAN1* genomic sequences were compared (81 and 64%; Table 2.3) was lower than had been observed when the corresponding coding sequences were compared (86 and 89%; Table 2.3). This lower level of identity resulted because the PCR-Bn1 and *RAN1* genomic sequences varied in both sequence and size (Figures 2.16 and 2.17). The PCR-Bn1 and *RAN1* predicted coding sequences were approximately the same size (Figures 2.14 and 2.15) and although they did vary in sequence, it was to a lesser extent than the corresponding genomic sequences. Since the only difference between the genomic and coding sequences is the presence of intron(s), the lower level of identity that is observed when the genomic sequences are compared can be attributed to differences in the introns. As a result, *Arabidopsis* sequence information cannot be expected to accurately predict the size of a *B. napus* PCR product if the primers span an intron-containing region and genomic DNA is used as the template. It is, however, likely that the *Arabidopsis* sequence information can be used to predict the size of a *B. napus* product that is produced from a coding sequence template or from a region of a genomic template that does not span any introns. Although the *Arabidopsis* and *B. napus* introns varied in both size and sequence, the introns disrupted the coding sequences in the same location. The *Arabidopsis* sequences might, therefore, be useful in predicting the location of *B. napus* introns.

Attempts were not made to clone, sequence, or analyze any of the remaining PCR products that had been produced with *B. napus* genomic DNA and the PCR-S/PCR-A

primer pair. Although the products that were not cloned may have originated from other, possibly novel $P_{1B}$-ATPases, it is also possible that these bands contain non-specific products or products from P-type ATPases that were not from the Type IB subfamily. In addition, the PCR-Bn1 and PCR-Bn2 products that had been cloned were likely amplified from genes that encode homologues of RAN1 and PAA1. Substrate specificity can be predicted for RAN1 and PAA1 with a relatively high level of confidence and it is likely that a similar prediction will be possible for the *B. napus* homologues of RAN1 and PAA1. When the PCR-Bn1 and PCR-Bn2 products were partially sequenced, *Arabidopsis* was still the only plant species for which full-length putative $P_{1B}$-ATPase sequences had been submitted to the NCBI databases and functional information was not yet available for any $P_{1B}$-ATPases from plants. The BnPCR1 and BnPCR2 products, if used in a cloning strategy to obtain full-length coding sequences, offered the possibility of reporting the first $P_{1B}$-ATPases from a plant species other than *Arabidopsis*. After cloning of the full-length sequence, the protein encoded by the full-length sequence could be examined using a *S. cerevisiae* complementation assay. Such a complementation assay could provide the first suggestion that functional $P_{1B}$-ATPases exist in plants and could be used to determine if the putative substrate specificity was accurately predicted by examining the amino acid sequence.

A decision was made to begin preparations to screen a *B. napus* cDNA library using probes based on the PCR-Bn1 and PCR-Bn2 products. The complete PCR-Bn1 and PCR-Bn2 fragments were not used to screen the amplified library as the nucleotides that encode the conserved amino acid sequences DKTGTL and GDGIND could potentially hybridize to clones that encode P-type ATPases that do not belong to the Type IB subfamily. In addition, blastn searches of the *Arabidopsis* "dbest" database highlighted additional nucleotide sequences within the PCR-Bn1 and PCR-Bn2 products that are commonly found in *Arabidopsis* coding sequences and might also be found in *B. napus* coding sequences.

### 2.3.2.2 Production of polymerase chain reaction-based probes and the screening of an amplified *Brassica napus* cDNA library

Probe1 and Probe2, which represent fragments of PCR-Bn1 and PCR-Bn2, were produced to screen the amplified *B. napus* cDNA library. These probes were produced by PCR, using the Probe1-S/Probe1-A and Probe2-S/Probe2-A primer pairs, and did not encode the DKTGTL and GDGIND motifs or contain the potentially common sequences that were discussed above.

The amplified *B. napus* cDNA library was first screened using Probe1. This probe yielded 16 primary positives from a total of 300 000 plaques. Nine of the primary positives were successfully followed through secondary screening, tertiary screening, and excision. Restriction digests of the excised plasmids revealed that the positives represented at least two inserts, one that was ~1000 bp in size and another that was ~2000 bp in size. Partial sequencing of select positives, which were named ~1000 bp cDNA and ~2000 bp cDNA, revealed that these positives were partial cDNAs that shared a high level of sequence similarity with the *RAN1* coding sequence from *Arabidopsis* (90% for ~1000 bp cDNA; 91 and 92% for ~2000 bp cDNA; Table 2.3). Both the ~1000 and ~2000 bp cDNAs appeared to be truncated, relative to the *RAN1* coding sequence, at the 5'-end. Sequencing of the ~2000 bp cDNA with the T3 primer revealed a stop codon in a position equivalent to the *RAN1* stop codon. A poly A tail was observed downstream of this stop codon. Although a full-length clone could potentially be isolated using Probe1, further screening with this probe was not done, as this screening would continually be complicated by the detection of partial clones (Figure 2.18).

The *RAN1* coding and genomic sequences were used to develop a strategy to produce a new probe that would, theoretically, provide a more efficient method, compared to Probe1, of identifying a full-length *B. napus* cDNA that encodes a RAN1 homologue. This new probe, which was named Probe3, was once again produced using a PCR-based strategy. The PCR primers were based on *Arabidopsis* sequences that were close to the 5'-end of the *RAN1* coding sequence. The location of these primers was such

that the resulting probe would not hybridize to any of the partial clones presently identified (Figure 2.18) and, as a result, would only detect other partial or full-length clones. The PCR primers, which were named Probe3-S and Probe3-A, were designed such that the *Arabidopsis* PCR product would not contain introns and the probability of a size difference between the *Arabidopsis* and *B. napus* PCR products would, consequently, be reduced.

The Probe3-S/Probe3-A primer pair was used in PCRs that contained either *Arabidopsis* or *B. napus* genomic DNA. An ~275 bp PCR product was amplified from the *Arabidopsis* genomic DNA. This product, which was very close in size to the 277 bp product that had been predicted using the *RAN1* sequences, was not cloned or sequenced. When the *B. napus* genomic DNA was used as a template three bands, which were ~235, ~255, and ~275 bp in size, were observed. The ~255 and ~275 bp products, which were successfully cloned and sequenced, were named Probe3-Bn1 and Probe 3-Bn2, respectively. These PCR products were found to share a high level of sequence identity with *RAN1* (77% for Probe3-Bn1; 82% for Probe3-Bn2; Table 2.3); a deletion, relative to *RAN1*, was, however, identified in both PCR products (Figure 2.19). The Probe3-Bn1 product contained a deletion that was 21 bp in length, while 3 bp were deleted from the same region of the Probe3-Bn2 product. Several nucleotides also varied between the two *B. napus* products (Figure 2.19). The results of this PCR, therefore, suggested that *B. napus* might contain at least two, and possibly three, *RAN1* homologues. The Probe3-Bn1 product was used to produce Probe3 and 2 100 000 plaques were screened with this new probe. No clear primary positives were identified and upon secondary screening all weak primary positives were eliminated.

Probe2, which had been produced from the PCR-Bn2 product and likely encoded a portion of a PAA1 homologue from *B. napus*, was similar to Probe1 in its potential to hybridize to partial clones. A second probe, which was named Probe4, was produced that would theoretically provide a more efficient method, compared to Probe2, of isolating a full-length *B. napus* cDNA that encodes a PAA1 homologue (Figure 2.2B). The production of this new probe was based on a strategy that was similar to the approach

used to produce Probe3. The *PAA1* coding and genomic sequences were used to design PCR primers that were based on relatively 5' sequences that did not span any introns. These primers were named Probe4-S and Probe4-A. The genomic sequence for *PAA1*, which was required to determine the location of introns, had been released as part of the *Arabidopsis* Genome Initiative (*PAA1*/AL035678/4490291).

The Probe4-S/Probe4-A primer pair was used in combination with *Arabidopsis* or *B. napus* genomic DNA. An ~290 bp PCR product was amplified from the *Arabidopsis* genomic DNA. This product, which was approximately the same size as the 289 bp product that had been predicted from the *RAN1* sequences, was not cloned or sequenced. When the *B. napus* genomic DNA was used as a template an ~280 bp fragment, which was slightly smaller than the *Arabidopsis* PCR product discussed above, was produced. Cloning and sequencing of a fragment from this ~280 bp band, which was named Probe4-Bn1, confirmed that the product shared a high level of sequence identity with the appropriate region of the *PAA1* sequences (68%; Table 2.3). The Probe4-Bn1 fragment was, therefore, used to produce Probe4. A total of 300 000 plaques were screened with Probe4 and no positives were identified. Probe2 and Probe4 were then used to screen 1 850 000 more plaques. Once again, no clear primary positives were identified and upon secondary screening all weak primary positives were eliminated.

Concurrent with the screening efforts, Hirayama *et al.* (1999) published a paper describing the cloning of a cDNA that encoded a putative $P_{1B}$-ATPase from *Arabidopsis*. Hirayama *et al.* (1999) also reported a number of experiments that suggested that the putative $P_{1B}$-ATPase that was encoded by this cDNA was able to transport copper. The cDNA that was cloned by Hirayama *et al.* (1999) corresponds to the T19K24.18 (*RAN1*) genomic locus from the *RAN1*/AC002342/2660661 genomic record. Since all of the cDNAs that had been isolated from the *B. napus* library had been found to share a high level of sequence identity with the *RAN1* coding sequence, a decision was made to briefly screen the amplified *B. napus* cDNA library for another putative $P_{1B}$-ATPase. The *Arabidopsis* sequence encoding the putative $P_{1B}$-ATPase HMA1 was selected to design two sets of PCR primers that were to be used to produce probes from *B. napus* genomic

DNA. The coding and amino acid sequences that had been predicted for the *HMA1* genomic entry that was identified in the original database search (*HMA1*/Z99707/2464848) had been updated (*HMA1*/Z99707/**4376087**) and this new coding sequence was used to design the primers. The HMA1 protein, and its homologues, are of interest since they might represent the first eukaryotic and plant $P_{1B}$-ATPases from the $Cd^{2+}$-ATPase division. It should, however, be noted that since the amino acid sequences that have been predicted for the *HMA1* genomic sequence appear to contain a variation of the CPx and GDGxNDxP motifs, it is possible that the HMA1 protein may transport a substrate that has yet to be reported for $P_{1B}$-ATPases. The first set of PCR primers, Probe5-S and Probe5-A, was based on a 5'-region of the *HMA1* sequence, while the second set of primers, Probe6-S and Probe6-A, was designed to amplify the region that encodes the SPC tripeptide.

Polymerase chain reactions (PCRs) were setup with the Probe5-S/Probe5-A primer pair and *Arabidopsis* or *B. napus* genomic DNA. An ~190 bp fragment was observed when the *Arabidopsis* genomic DNA was used as the template. The size of this product was similar to that of the 188 bp fragment that had been predicted using the *HMA1* sequence. A much larger product, ~830 bp, was produced when *B. napus* genomic DNA was used as a template. Partial sequencing of this ~830 bp product revealed that it had been amplified using the Probe5-A primer as both the sense and antisense primer and that this product did not share a high level of sequence similarity with *HMA1* or any other P-type ATPase (data not shown). Cloning and sequencing of a fragment from the ~190 bp band observed in the *Arabidopsis* PCR confirmed that the Probe5-S/Probe5-A primer pair was functional and produced the expected product, which was named Probe5-At1, when *Arabidopsis* genomic DNA was supplied as the template (Table 2.3).

Polymerase chain reactions (PCRs) were also setup for the Probe6-S/Probe6-A primer pair and either *Arabidopsis* or *B. napus* genomic DNA. When the *Arabidopsis* genomic DNA was used as a template, an ~375 bp band, which is close in size to the 373 bp fragment that was predicted from the *HMA1* genomic sequence, was produced.

Cloning and sequencing of a fragment from this ~375 bp band revealed that the fragment, which was named Probe6-At1, had likely originated from the *HMA1* genomic sequence (Table 2.3). Translation of the appropriate region of the Probe6-At1 product confirmed that the SPC tripeptide was found in place of the CPx motif. An ~375 bp band was produced when the Probe6-S/Probe6-A primer pair was used in combination with *B. napus* genomic DNA. A fragment from this band was cloned and sequenced. This fragment, which was named Probe6-Bn1, was approximately the same size as the Probe6-At1 product and shared a high level of sequence identity with *HMA1* (93% identity when coding sequences were compared; Table 2.3). Interestingly, the Probe6-Bn1 product encoded a SQC tripeptide instead of the CPx motif that is observed in most putative $P_{1B}$-ATPases or the SPC tripeptide that is observed in HMA1. Probe6 was produced from the Probe6-Bn1 product and used to screen a total of 700 000 plaques. There were no clear primary positives identified and upon secondary screening any weak primary positives were eliminated.

#### 2.3.2.3 Isolation of the 5'-coding sequence that is missing from the ~2000 bp library clone by reverse transcriptase-polymerase chain reaction and production of a full-length cDNA

The largest clone that had been isolated from the amplified *B. napus* cDNA library was an ~2000 bp cDNA that shared a high level of sequence identity with the *RAN1* coding sequence and appeared to be truncated at the 5'-end of the coding sequence. The *RAN1* and ~2000 bp cDNA sequences were, consequently, examined to determine if a strategy could be developed to isolate the missing *B. napus* 5'-coding sequence and to ligate it to the ~2000 bp cDNA. A restriction analysis, using the DNAMAN program, of the *Arabidopsis RAN1* coding sequence revealed a unique *SacI* site downstream of the region where the *RAN1* coding sequence and the ~2000 bp cDNA began to overlap. The available sequence information for the ~2000 bp cDNA suggested that this *SacI* site was also present in the ~2000 bp cDNA. Restriction digests of the ~2000 bp cDNA confirmed that the *SacI* site was present and unique in the ~2000 bp cDNA.

A strategy was, therefore, developed to amplify a RT-PCR product that would contain the missing *B. napus* 5'-coding sequence and the *SacI* site. The *SacI* site, if unique in the RT-PCR product containing the 5'-coding sequence, could be used to join the missing *B. napus* 5'-coding sequence to the ~2000 bp cDNA. A sense primer, RT1-S, was designed from the 5'-untranslated region of the *RAN1* mRNA/CDS sequence (AF091112/4760379). A compatible antisense primer, RT1/2-A, was designed so that the missing 5'-coding sequence and ~478 bp of sequence that overlapped with the ~2000 bp cDNA and contained the *SacI* restriction site would be produced. The overlapping region allowed nucleotide sequences from the ~2000 bp cDNA and the RT-PCR product to be compared. The RT-PCRs were setup with HotStarTaq™ DNA Polymerase to increase specificity. All RT-PCRs that were set up with the first sense primer, RT1-S, were unsuccessful and a second sense primer was, consequently, designed. This second primer, which was named RT2-S, was designed to be compatible with RT1/2-A and was based on the 5'-untranslated region and several of the initial nucleotides found in the *RAN1* mRNA/CDS (AF091112/4760379).

A series of RT-PCRs were setup in which first strand cDNA was produced with the Probe1-A, RT3-A, or RT1/2-A primers and the final products were amplified using the RT2-S and RT1/2-A primers. A faint band of ~1490 bp, which was ~160 bp larger than the product that had been predicted from the *Arabidopsis RAN1* mRNA/CDS record (AF091112/4760379), was observed in the RT-PCRs that had used the Probe1-A and RT3-A primers for first strand cDNA synthesis (Figure 2.20). Cloning and partial sequencing of a fragment from this ~1490 bp band revealed that this band probably resulted from contamination of the RT-PCRs with *B. napus* genomic DNA, as the cloned fragment appeared to contain introns (data not shown). A ~1310 bp product, which was similar in size to the 1333 bp fragment that had been predicted from the *RAN1* mRNA/CDS record (AF091112/4760379), was also observed in all of the RT-PCRs (Figure 2.20). Restriction digests of fragments that had been cloned from these ~1310 bp bands confirmed that a unique *SacI* site was present in the RT-PCR products and allowed the identification of fragments that had been cloned in the appropriate orientation for ligation to the partial library clone. A *RsaI* restriction analysis of the fragments that had

been cloned in the desired orientation suggested that there were at least two products within the ~1310 bp bands. Partial sequencing of a representative from the first group revealed that this RT-PCR product, which was named RT-Bn1, shared a high level of sequence identity with the *RAN1* coding sequence (79 and 92%; Table 2.3). A 21 bp deletion, relative to the *RAN1* coding sequence, was observed within RT-Bn1. This same deletion had been observed in the Probe3-Bn1 product. A comparison of the partial sequencing results for RT-Bn1 and the ~2000 bp cDNA revealed 9 differences in the 339 bp of overlapping sequence information that was available (Figure 2.21). Partial sequencing of a representative from the second group of ~1310 bp RT-PCR products, which was named RT-Bn2, revealed a high level of sequence identity between this fragment and the *RAN1* coding sequence (82 and 93%; Table 2.3). A 3 bp deletion, relative to the *RAN1* coding sequence, was observed in the RT-Bn2 product. The Probe3-Bn2 product also contained this deletion. When the partial sequence from the RT-Bn2 product was compared to the ~2000 bp library clone there were no differences observed in the 339 bp of overlapping sequence information that was available (Figure 2.22). The identification of Probe3-Bn1 and Probe3-Bn2, two different PCR products from the Probe3-S and Probe3-A primers, suggested that *B. napus* genome might encode at least two RAN1 homologues. The RT-Bn1 and RT-Bn2 products provide additional support for this suggestion. The above results also indicate that the second group of RT-PCR products, which include RT-Bn2, likely include the 5'-coding sequence that is missing from the ~2000 bp cDNA and the RT-Bn2 product was, therefore, used in the ligation reaction. The appropriate fragments from the RT-Bn2 product and the ~2000 bp cDNA were successfully ligated into the pBluescript SK- vector to form a full-length *B. napus* cDNA. All of the vectors containing the appropriately assembled ligation products were, however, maintained at low levels in *E. coli*.

#### 2.3.2.4 Sequencing and analysis of the full-length *Brassica napus* cDNA

Sequencing of a region of the full-length *B. napus* cDNA that had been derived from the RT-Bn2 product revealed a single base pair deletion, which resulted in a frame

shift and premature stop codons (data not shown). This deletion was verified by a second sequencing reaction. The remaining RT-PCR products that had been cloned and found to display the same *RsaI* restriction pattern as the RT-Bn2 product were sequenced in this region. A RT-PCR product, which was named RT-Bn3, was identified that did not contain the single base pair deletion. When the partial sequence from the RT-Bn3 product was compared to the ~2000 bp library clone a single difference was observed in the 339 bp of overlapping sequence information that was available (Figure 2.23). This nucleotide variation results in an amino acid change from aspartate (D) to Glycine (G). This aspartate (D), which is equivalent to residue 337 of RAN1, is not conserved amongst the $P_{1B}$-ATPases and a decision was made to ligate the RT-Bn3 product to the ~2000 bp cDNA.

Sequencing and translation of the full-length *B. napus* cDNA that contained a portion of RT-Bn3 and a portion of the ~2000 bp cDNA confirmed that it was free of frameshifts and inappropriate stop codons. This *B. napus* cDNA sequence has been entered into the GenBank database as a mRNA/CDS record (AY045772/15636780). The predicted amino acid sequence is also available as a protein record (AAL02122/15636781). The amino acid sequence that is encoded by the *B. napus* cDNA contains all of the P-type (Section 1.1) and Type IB (Section 1.4) ATPase motifs that have been discussed previously. The N-terminal metal-binding region consists of two copies of the CxxC motif. The hydrophobicity profile that is produced from this *B. napus* amino acid sequence (Figure 2.24) is most similar to that of other $P_1$-ATPases/$P_{1B}$-ATPases (Figure 2.12A). The N-terminus has a relatively large number of hydrophobic regions, while the C-terminus contains a relatively low number of hydrophobic regions. These results suggest that the *B. napus* cDNA that was cloned in this study encodes a putative $P_{1B}$-ATPase.

A blastp search revealed that the amino acid sequence encoded by the *B. napus* cDNA is similar to RAN1. When Hirayama *et al.* (1999) cloned the *RAN1 Arabidopsis* cDNA (AF091112/4760379) it was determined that the coding and protein sequences that had been predicted for the original *RAN1* genomic entry (*RAN1*/AC002342/2660661

genomic locus T19K24.18) were incorrect. The coding and amino acid sequences in this genomic entry were adjusted; a new GI identifier was not, however, assigned. The original protein entry (RAN1/AAC79141/2660670) was also altered and, in this case, a new GI identifier (RAN1/AAC79141/**6850337**) was assigned. The amino acid sequence that was predicted from the *B. napus* cDNA sequence was 91% identical to the updated RAN1 amino acid sequence (RAN1/AAC79141/**6850337**). Four regions were identified where the *B. napus* amino acid sequence displayed a deletion or addition relative to the updated RAN1 amino acid sequence (RAN1/AAC79141/**6850337;** Figure 2.25). The full-length *B. napus* coding sequence was 88% identical to the *RAN1* coding sequence that was reported after the work of Hirayama *et al.* (1999). These results suggest that the *B. napus* cDNA encodes a RAN1 homologue, while the PCR and RT-PCR results (Sections 2.3.2.2 and 2.3.2.3) suggest the cloned *B. napus* cDNA may encode just one of two or more RAN1 homologues that are found in *B. napus*.

In addition to RAN1, the ATP7B and ATP7A proteins were identified as top blastp matches when the *B. napus* amino acid sequence was used as a query. The *B. napus* amino acid sequence was ~28% identical to ATP7B and ~27% identical to ATP7A. It is important to note that these identity levels were reduced by the extended N-terminus of the ATP7A and ATP7B proteins and the fact that P-type ATPases only contain dispersed regions of high sequence identity (Axelsen and Palmgren, 1998). Evidence has been presented to suggest that the RAN1, ATP7A, and ATP7B proteins are able to transport copper (Hung *et al.*, 1997; Forbes and Cox; 1998; Iida *et al.*, 1998; Payne and Gitlin, 1998; Hirayama *et al.*, 1999; Forbes, 2000) and copper transport has been directly demonstrated for ATP7A (Voskoboinik *et al.*, 1998). Since the P-type ATPases group according to substrate specificity (Axelsen and Palmgren, 1998), the *B. napus* cDNA is likely to encode a protein that is capable of transporting copper. This suggestion is further supported by the presence of leucine residues (L) 21 amino acids C-terminal of the two CxxC motifs that are found in the amino acid sequence that was predicted from the *B. napus* cDNA. The amino acid residue that is located 21 amino acids C-terminal of the CxxC motif(s) may modulate metal-binding affinity/specificity (Gitschier *et al.*, 1998). A leucine (L) residue is present 21 amino acids C-terminal of the

CxxC motif(s) that are found in many $Cu^{2+}$-ATPases, including CopA, Ccc2p, ATP7A, and ATP7B (Gitschier *et al.*, 1998).

## 2.4 Discussion

The results of the literature and database searches provided support for the existence of plant genes that encode putative $P_{1B}$-ATPases. Four full-length sequences, which have now been named *PAA1*, *RAN1*, *HMA1*, and *HMA4*, were identified. All of these sequences were from *Arabidopsis*. The PAA1 and RAN1 proteins are more similar to the $P_{1B}$-ATPases that have been studied in detail than are the HMA1 and HMA4 proteins. The PAA1 and RAN1 predicted amino acid sequences contained all of the P-type and Type IB motifs that are commonly found in $P_{1B}$-ATPases. One P-type motif and one Type IB motif were altered in the HMA1 sequence. The HMA4 amino acid sequence contained all of the P-type motifs, but was missing one of the three Type IB motifs. The hydrophobicity profiles that were predicted for the PAA1, RAN1, HMA1, and HMA4 amino acid sequences were similar to those predicted for other $P_1$-ATPases/$P_{1B}$-ATPases. Expressed sequence tags (ESTs) or cDNA/mRNA entries that corresponded to the *PAA1*, *RAN1*, *HMA1*, and *HMA4* coding sequences were identified, suggesting that these genes are transcribed and supporting all or part of the protein predictions. Three fragments that likely correspond to putative $P_{1B}$-ATPases from *Glycine max* (soybean), *Gossypium hirsutum* (cotton), and *Oryza sativa* (rice) were also identified, suggesting that genes encoding putative $P_{1B}$-ATPases might also be found and transcribed in plant species other than *Arabidopsis*.

Sequence databases are constantly growing and records have now been added to the databases that provide additional information about PAA1, RAN1, HMA1, and HMA4. The genomic sequence (*PAA1*/AL035678/4490291 genomic locus F17M5.280) that corresponds to the *PAA1* mRNA/CDS is now available. A protein record (HMA4/AAD12041/4210504) for the HMA4 amino acid sequence is also now available. Hirayama *et al.* (1999), who reported the cloning of a cDNA that encodes RAN1, have submitted a *RAN1* mRNA/CDS entry (AF091112/4760379). Although the cloning of

cDNAs that encoded HMA1 and HMA4 has not been reported in the literature, the addition of *HMA1* (AJ400906/7710953) and *HMA4* (AJ297264/11990206) mRNA/CDS entries suggest that cDNAs that encode these proteins have now been cloned. Additional information is not yet available for the *Glycine max* (soybean; AF019115/2565258), *Gossypium hirsutum* (cotton; AI054702/3325816), and *Oryza sativa* (rice; D42886/3107148) partial sequences.

Some of the entries that were identified in the initial database searches have now been revised. The cloning of a cDNA encoding RAN1 by Hirayama *et al.* (1999) revealed that the coding and protein sequences that had been predicted for the original *RAN1* genomic entry (*RAN1*/AC002342/2660661 genomic locus T19K24.18) were incorrect. The coding and protein sequences that are reported in the *RAN1* genomic entry have, consequently, been adjusted; however, a new GI identifier was not assigned to reflect these changes. The original RAN1 protein entry (RAN1/AAC79141/2660670) was also altered and, in this case, a new GI identifier (RAN1/AAC79141/**6850337**) was assigned. The predicted coding and protein sequences that correspond to the original *HMA1* genomic entry (*HMA1*/Z99707/2464848) have also been altered. The genomic and protein entries were each assigned new accession numbers and/or GI identifiers (*HMA1*/Z99707/**4376087** and HMA1/**CAB16773/4006855**) to reflect this change. The *HMA1* mRNA/CDS entry (AJ400906/7710953) supports the revised HMA1 protein prediction (HMA1/**CAB16773/4006855**). Although the original *HMA4* genomic entry (*HMA4*/AC002392/3176701 genomic locus T20K24.12) has been assigned a new GI identifier (*HMA4*/AC002392/**6598371** genomic locus T20K24.12), this new GI identifier does not correspond to a change in the coding or protein sequences that were originally predicted for the *HMA4* gene. The *HMA4* mRNA/CDS entry (AJ297264/11990206) supports the protein prediction that was reported for the both of the *HMA4* genomic entries.

The revised RAN1 (RAN1/AAC79141/**6850337**) and HMA1 (HMA1/**CAB16773/4006855**) proteins contained all of the P-type and Type IB motifs that were identified in the original RAN1 (RAN1/AAC79141/2660670) and HMA1

(HMA1/2464854/2464854) amino acid sequences. Although the changes in the RAN1 (RAN1/AAC79141/**6850337**) and HMA1 (HMA1/**CAB16773/4006855**) amino acid sequences did alter the hydrophobicity profile (data not shown), the new hydrophobicity profiles were still more similar to the profiles of other $P_1$-ATPases/$P_{1B}$-ATPases than the profiles of $P_2$- or $P_3$-ATPases.

Database searches were performed on a regular basis throughout this study. Five additional full-length putative $P_{1B}$-ATPases that were from plant species were identified in these database searches. The first four putative $P_{1B}$-ATPases were from *Arabidopsis* and contained both genomic (*HMA2*/AL078464/4938473 genomic locus F6G3.140, *HMA3*/AL078464/4938473 genomic locus F6G3.150, *HMA5*/AC008047/6623889 genomic locus F2K11.18, and "*HMA6*"/AL589883/13374848 AGI name At5g21930) and protein (HMA2/CAB43846/4938487, HMA3/CAB43847/4938488, HMA5/AAF19707/6633848, and "HMA6"/CAC34486/13374852) entries. The "*HMA6*"/"HMA6" gene/protein name was assigned to the AL589883/13374848 AGI name At5g21930 and CAC34486/13374852 records in conjunction with this study, as these records have not been discussed in the literature or assigned a gene/protein name. Although mRNA/CDS entries are not available for these latest *Arabidopsis* sequences, ESTs (data not shown) confirm regions of the HMA2 and HMA6 protein predictions. The fifth full-length sequence was entered in conjunction with this study and includes the sequence of the *B. napus* cDNA (AY045772/15636780) that was cloned in this study. This *B. napus* cDNA encodes the first full-length putative $P_{1B}$-ATPase (AAL02122/15636781) in the NCBI databases that is from a plant other than *Arabidopsis* and, therefore, provides the most compelling evidence yet suggesting that $P_{1B}$-ATPases are also found in plant species other than *Arabidopsis*.

The amino acid sequences that are encoded by *HMA5*, *HMA6*, and the *B. napus* cDNA contain all of the P-type (Section 1.1) and Type IB (Section 1.4) motifs that have been discussed previously. The N-terminal metal-binding region of HMA5 contains one copy of the CxxC motif, while HMA6 and the *B. napus* proteins each contain two copies of the CxxC motif. The hydrophobicity profiles predicted for the HMA5, HMA6 (data

not shown), and *B. napus* (Figure 2.24) amino acid sequences are more similar to the profiles of $P_1$-ATPases/$P_{1B}$-ATPases than the $P_2$- or $P_3$-ATPases. The HMA2 and HMA3 amino acid sequences contain all of the previously outlined P-type motifs (Section 1.1) and two of the three Type IB motifs (Section 1.4). The amino acid sequences that were predicted from the *HMA2* and *HMA3* coding sequences did not contain an N-terminal metal-binding region. The C-terminus of HMA2 contains six cysteine (C) pairs and a histidine-rich (H-rich) region, while the C-terminus of HMA3 contained two cysteine (C) pairs. As was proposed for HMA4 (personal observations; Williams *et al.*, 2000), the histidine-rich (H-rich) region and/or the cysteine (C) pairs could potentially act as metal-binding domains. The HMA2, HMA3, and HMA4 proteins might, therefore, be the first of a group of $P_{1B}$-ATPases that contain C-terminal metal-binding regions. The HMA2 and HMA3 hydrophobicity profiles (data not shown) were more similar to the profiles of $P_1$-ATPases/$P_{1B}$-ATPases than $P_2$- or $P_3$-ATPases.

Expressed sequence tags (ESTs) that encode amino acid sequences that share a high level of sequence identity with $P_{1B}$-ATPases and contain at least one Type IB motif (Section 1.4) can now be identified from numerous plant species. Included amongst these plant species are economically important species such as *Triticum aestivum* (bread wheat; E405529/9364997), *Hordeum vulgare* (barley; BE215627/8903155), *Populus balsamifera* subsp. *trichocarpa* (poplar; AI166493/3857778), *Solanum tuberosum* (potato; BI176444/14642255), *Beta vulgaris* (sugar beet; BE590297/9843345), and *Lycopersicon esculentum* (tomato; AW037991/5896745). These ESTs suggest that $P_{1B}$-ATPases are found in diverse plant species and offer information that might aid in the cloning of the corresponding cDNAs.

The $P_{1B}$-ATPase subfamily can be divided into at least two groups; the first group contains all of the $Cu^{2+}$-ATPases and the second group contains the $Cd^{2+}$-ATPases (Axelsen and Palmgren, 1998). The PAA1, RAN1 (original and revised), HMA5, HMA6, and *B. napus* proteins are most similar to the $Cu^{2+}$-ATPases and it is probable that these proteins are involved in copper transport. The HMA1 (original and revised), HMA2, HMA3, and HMA4 proteins, which are most similar to the $Cd^{2+}$-ATPases, may

be involved in cadmium transport. The HMA1, HMA2, HMA3, and HMA4 amino acid sequences are, however, the first, or amongst the first, eukaryotic putative $P_{1B}$-ATPases to share a higher level of similarity with the $Cd^{2+}$-ATPases than the $Cu^{2+}$-ATPases. The HMA1, HMA2, HMA3, and HMA4 proteins also each contain one putative Type IB motif that differs from the corresponding Type IB motif in $P_{1B}$-ATPases that have been studied in detail. It is, therefore, possible that the HMA1, HMA2, HMA3, and/or HMA4 proteins may display a substrate specificity that has not yet been reported for the $P_{1B}$-ATPases.

When this project was initiated the amount of information that was available for $P_{1B}$-ATPases from plants was limited. Four full-length plant sequences, which have now been named *PAA1*, *RAN1*, *HMA1*, and *HMA4*, encoding putative $P_{1B}$-ATPases were reported in the literature and/or sequence databases. Although a cDNA encoding PAA1 had been cloned (Tabata *et al.*, 1997), there was no experimental evidence to suggest that any of the putative $P_{1B}$-ATPases were functional or to investigate their substrate specificity. During the progress of this study, functional information was published for the *Arabidopsis* putative $P_{1B}$-ATPase RAN1 (Hirayama *et al.*, 1999; Woeste and Kieber, 2000). A number of reviews (Palmgren and Harper, 1999; Williams *et al*, 2000) also briefly mention studies focussing on a putative $P_{1B}$-ATPase that has been named AMA1 for *Arabidopsis* heavy metal ATPase 1. It is unclear which, if any, of the putative *Arabidopsis* $P_{1B}$-ATPases that are found in the databases are equivalent to AMA1. The cloning of a *B. napus* cDNA that encodes a putative $P_{1B}$-ATPase, therefore, offered the opportunity to add to the limited amount of functional data that was available for putative plant $P_{1B}$-ATPases.

The amino acid sequence that was predicted from the *B. napus* cDNA isolated in the current study groups with the $Cu^{2+}$-ATPases. The predicted *B. napus* amino acid sequence is most similar to the updated RAN1 amino acid sequence (RAN1/AAC79141/**6850337;** ~91% identical). The *B. napus* amino acid sequence is also similar to the ATP7B (~28% identical) and ATP7A (~27% identical) proteins, which represent top matches when the *B. napus* amino acid sequence is used as a blastp query.

*Saccharomyces cerevisiae* strains that have had the *CCC2* gene disrupted have become an important tool for studying putative $Cu^{2+}$-ATPases from other organisms. The ability of a putative $Cu^{2+}$-ATPase to rescue the *ccc2* mutant has been interpreted as evidence suggesting that the protein transports copper (Hung *et al.*, 1997; Sambongi *et al.*, 1997; Forbes and Cox, 1998; Iida *et al.*, 1998; Payne and Gitlin, 1998; Hirayama *et al.*, 1999; Forbes, 2000). The ATP7A and ATP7B proteins, which are both ~25% identical to Ccc2p, are amongst the proteins that have been shown to complement the *ccc2* mutant (Hung *et al.*, 1997; Forbes and Cox, 1998; Iida *et al.*, 1998; Payne and Gitlin, 1998; Forbes, 2000). The amino acid sequence that is predicted for the *B. napus* cDNA is ~34% identical to Ccc2p. A *ccc2* complementation assay will, consequently, be used to determine if the *B. napus* cDNA that was cloned in this study encodes a functional copper-transporting protein.

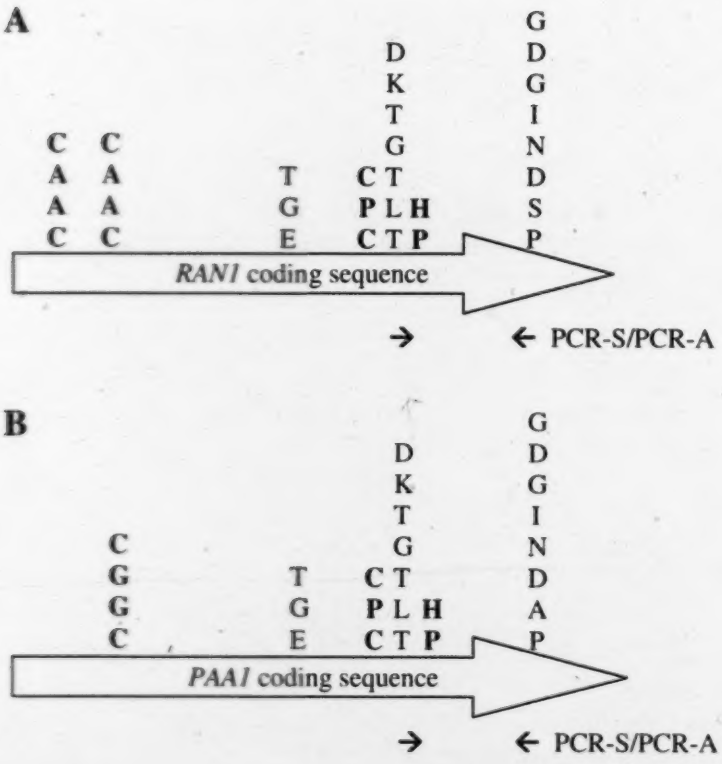


**Figure 2.1 The position of the PCR-S and PCR-A primers relative to the *RAN1* (A) and *PAA1* (B) coding sequences.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *RAN1* and *PAA1* coding sequences are indicated above the arrow that represents the appropriate coding sequence. The partially degenerate primers PCR-S and PCR-A were based on the *RAN1* and *PAA1* nucleotides that encode the DKTGTL and GDGIND amino acids.

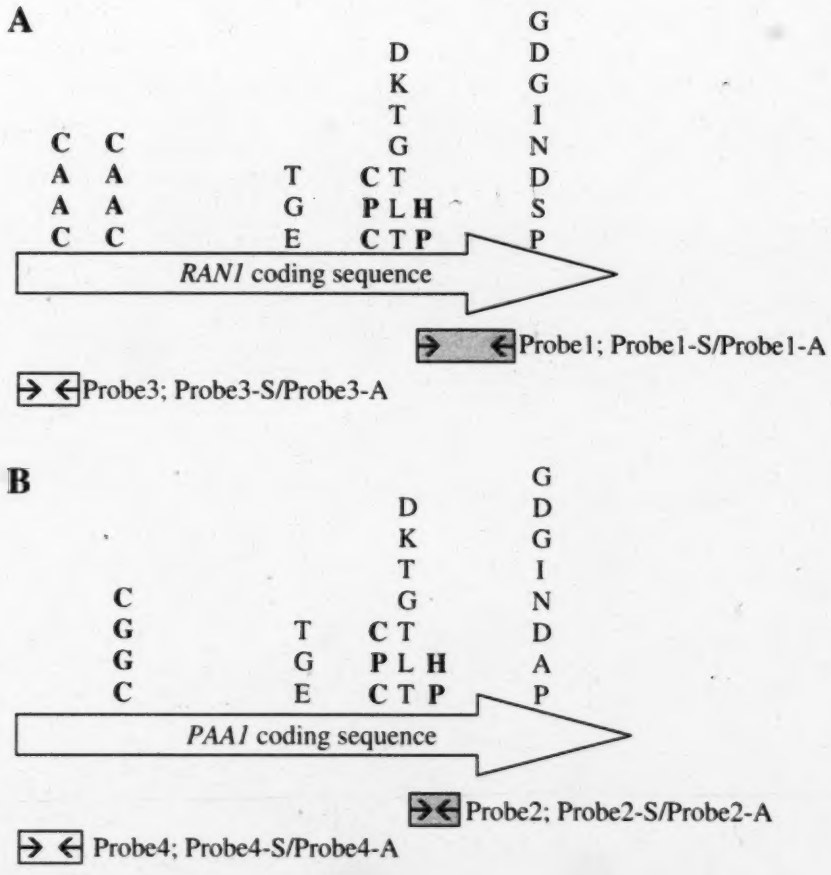


**Figure 2.2 The position of the primers and probes that were produced in preparation to screen the amplified *B. napus* library for cDNAs that encode RAN1 (A) or PAA1 (B) homologues.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *RAN1* and *PAA1* coding sequences are indicated above the arrow that represents the appropriate coding sequence. **A.** Probe1 and Probe3 are positioned relative to the *RAN1* coding sequence. Probe1 was derived from a PCR product that was amplified using the Probe1-S/Probe1-A primer pair and the PCR-Bn1 product. The rectangle representing Probe1 has been shaded to indicate that Probe1 contains an intron that is not reflected in its alignment with the *RAN1* coding sequence. Probe3 was derived from the Probe3-Bn1 PCR product, which was amplified using the Probe3-S/Probe3-A primer pair and *B. napus* genomic DNA. **B.** Probe2 and Probe4 are positioned relative to the *PAA1* coding sequence. Probe2 was derived from a PCR product that was amplified using the Probe2-S/Probe2-A primer pair and the PCR-Bn2 product. The rectangle representing Probe2 has been shaded to indicate that Probe2 contains an intron that is not reflected in its alignment with the *PAA1* coding sequence. Probe4 was derived from the Probe4-Bn1 PCR product, which was amplified using the Probe4-S/Probe4-A primer pair and *B. napus* genomic DNA.

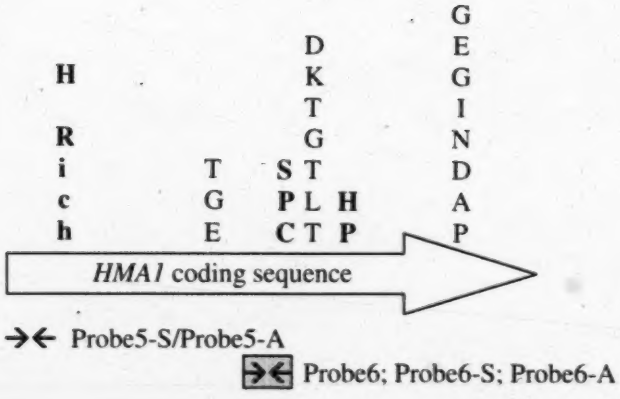


**Figure 2.3 The position of the primers and a probe that were produced in preparation to screen the amplified *B. napus* library for cDNA(s) that encode a HMA1 homologue.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *HMA1* coding sequence are indicated above the arrow that represents the coding sequence. Probe6 is positioned relative to the revised *HMA1* coding sequence (*HMA1*/Z99707/**4376087**). Probe6 was derived from the Probe6-Bn1 PCR product, which was amplified using the Probe6-S/Probe6-A primer pair and *B. napus* genomic DNA. The rectangle representing Probe6 has been shaded to indicate that Probe6 contains an intron that is not reflected in its alignment with the revised *HMA1* coding sequence. The Probe5-S/Probe5-A primer pair, which did not result in the production of a probe that could be used to screen the *B. napus* cDNA library, is also positioned relative to the *HMA1* coding sequence.

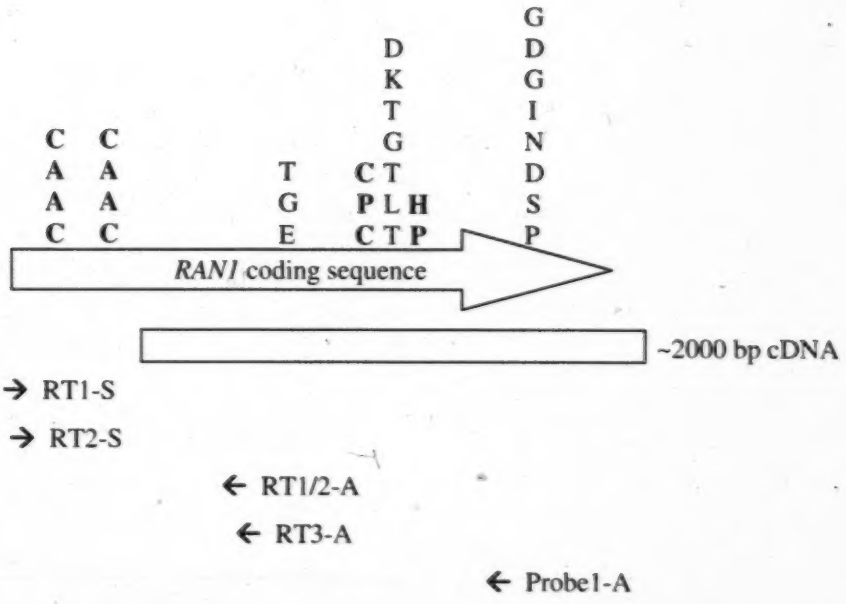


**Figure 2.4 The position of the primers used in the RT-PCR strategy that was intended to amplify the 5'-coding region that corresponds to the ~2000 bp cDNA.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *RAN1* coding sequence are indicated above the arrow that represents the coding sequence. The primers are positioned relative to the *RAN1* coding sequence and the ~2000 bp partial library clone. The 5'-untranslated region of the *RAN1* mRNA/CDS (AF091112/4760379) sequence was used to design the RT1-S primer. The RT2-S primer was based on the 5'-untranslated region and several of the initial nucleotides found in the *RAN1* coding sequence. The RT1/2-A primer was designed to be compatible with both the RT1-S and RT2-S primers. This anitsense primer was based on the sequencing information that had been obtained from a preliminary sequencing reaction of the ~2000 bp cDNA. The RT3-A and Probe1-A primers were based on sequencing information from the ~2000 bp cDNA and the PCR-Bn1 product, respectively.

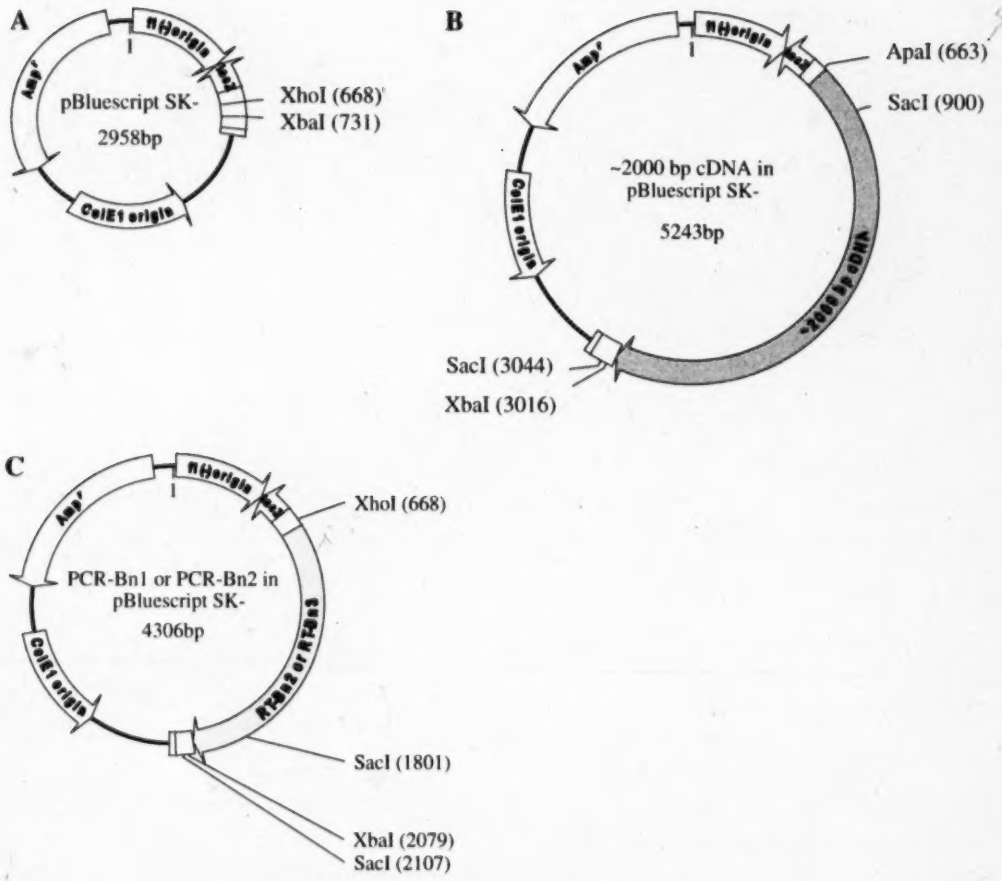


**Figure 2.5 Plasmids that were used in the ligation reaction to produce a full-length *B. napus* cDNA. A.** The pBluescript SK- vector was digested with *XhoI* and *XbaI*. The ~2895 bp fragment was gel isolated and used in the ligation reaction. **B.** The ~2000 bp cDNA that was isolated from the amplified *B. napus* library was liberated from the pBluescript SK- vector by a *ApaI* / *XbaI* digest. The ~2350 bp fragment was then gel isolated and digested with *SacI*. The ~2115 bp fragment from this *SacI* digest was gel isolated and used in the ligation reaction. **C.** The RT-Bn2 or RT-Bn3 RT-PCR products were liberated from the pBluescript SK- vector by a *XhoI* / *XbaI* digest. The ~1410 bp fragment was then gel isolated and digested with *SacI*. The ~1130 bp fragment from this *SacI* digest was then gel isolated and used in the ligation reaction.

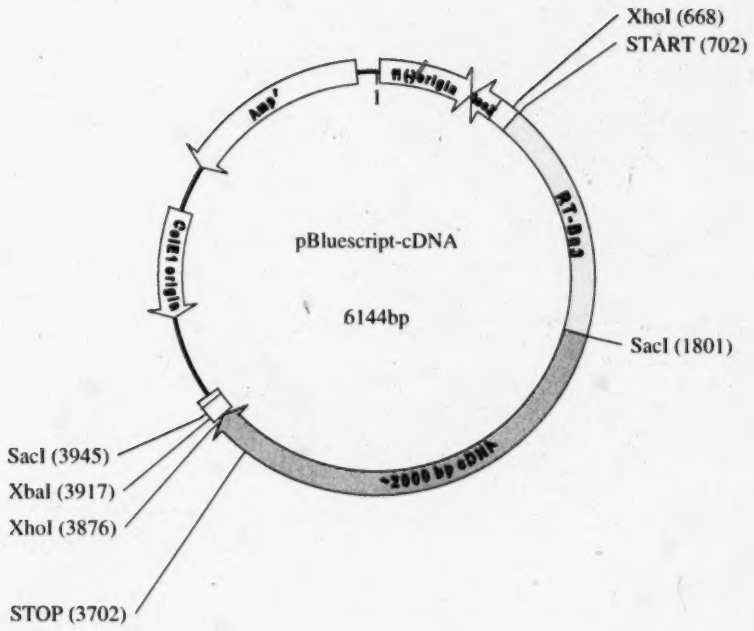


**Figure 2.6 The full-length *B. napus* cDNA in the pBluescript SK- plasmid (pBluescript-cDNA).** The pBluescript-cDNA was created by cloning the RT-Bn3 (or RT-Bn2) RT-PCR fragment with *XhoI* / *SacI* ends (light grey) and the ~2000 bp cDNA fragment with *SacI* / *XbaI* ends (dark grey) into a *XhoI* / *XbaI* digested pBluescript SK- vector (white).

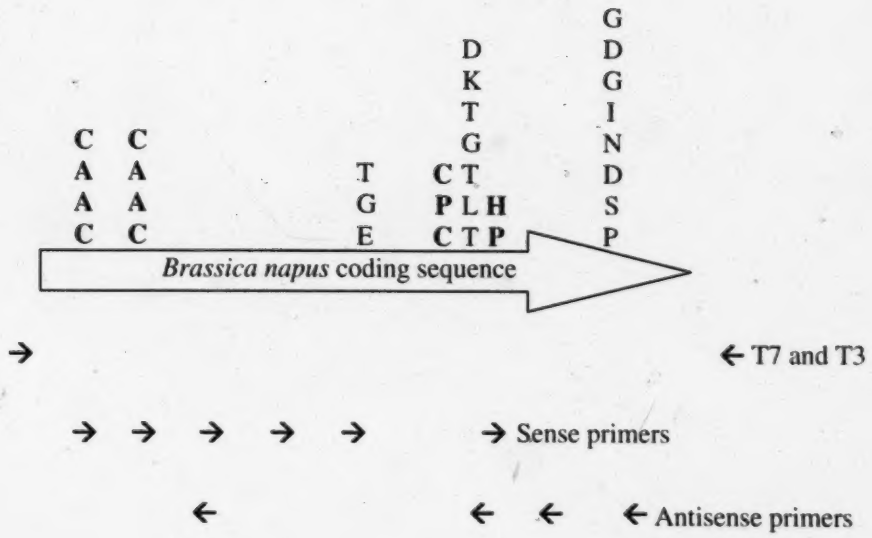


**Figure 2.7 The position of primers that were used to sequence the *B. napus* cDNA relative to the *B. napus* coding sequence.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *B. napus* cDNA are indicated above the arrow that represents the coding sequence. The T7 and T3 primers are in the first row. The second row of primers includes all of the sense primers (from left to right; Seq1-S, Seq2-S, Seq4-S, Seq5-S, Seq6-S, and Probe1-S). The final row of primers includes all of the antisense primers (from left to right; Seq3-A, Seq7-A, Seq8-A, and Seq9-A). With the exception of the T7 and T3 primers, these primers were based on sequencing results from the full-length *B. napus* cDNA or the PCR-Bn1 product.

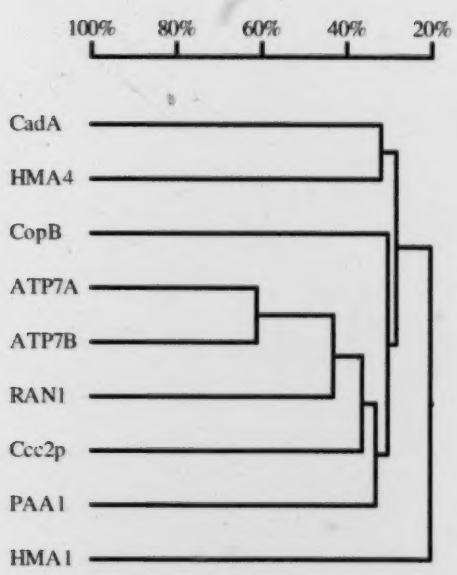


**Figure 2.8 Phylogenetic tree.** The CadA (AAB59154/150719), CopB (AAA61836/290643), ATP7A (AAA35580/179253), ATP7B (AAB52902/1947035), Ccc2p (AAC37425/538515), PAA1 (BAA23769/2668492), RAN1(AAC79141/2660670), HMA1(2464854/2464854), and HMA4 (amino acid sequence corresponding to genomic locus T20K24.12 from AC002392/3176701) sequences were aligned using the default settings of the optimal alignment option that is available as part of the DNAMAN multiple alignment tool. The phylogenetic tree was created using the homology tree feature that is available from the output button of the multiple alignment results.

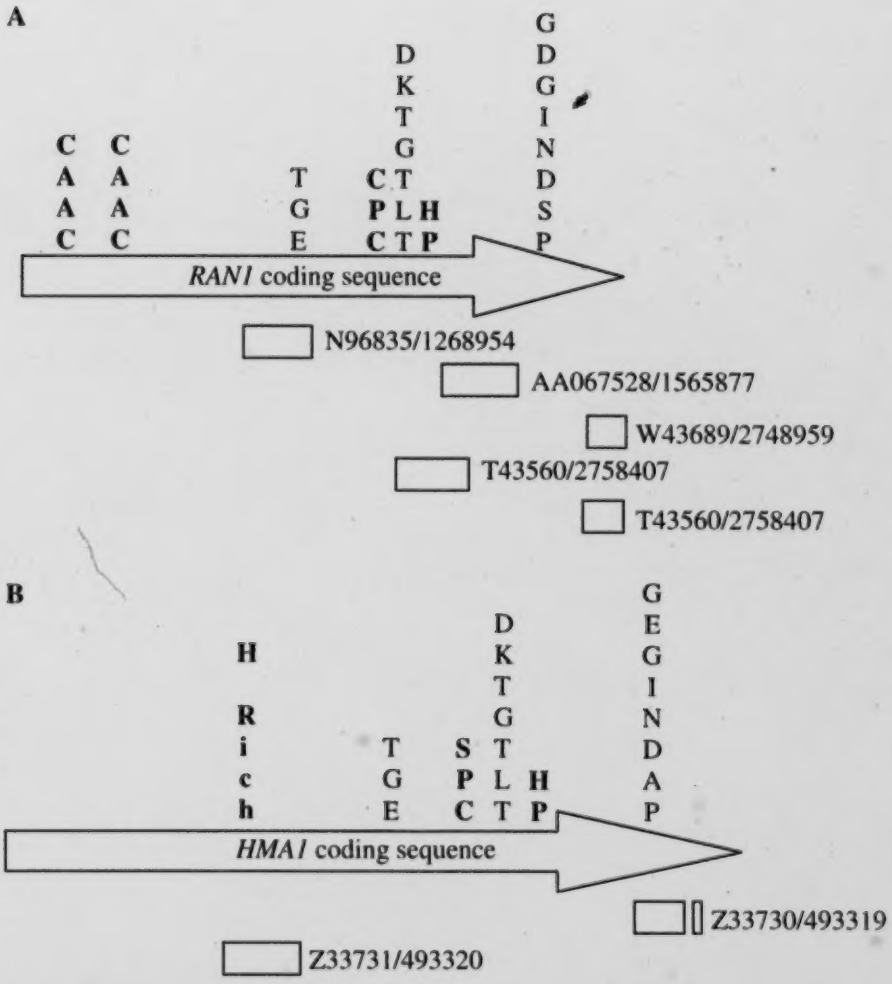


**Figure 2.9 Expressed sequence tags (ESTs) that correspond to the *RAN1* (A) and *HMA1* (B) coding sequences.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *RAN1* and *HMA1* coding sequences are indicated above the arrow that represents the appropriate coding sequence. The ESTs, which were identified by a variety of BLAST searches (Sections 2.2.1.2 and 2.2.1.3.3), are positioned relative to the appropriate coding sequence.

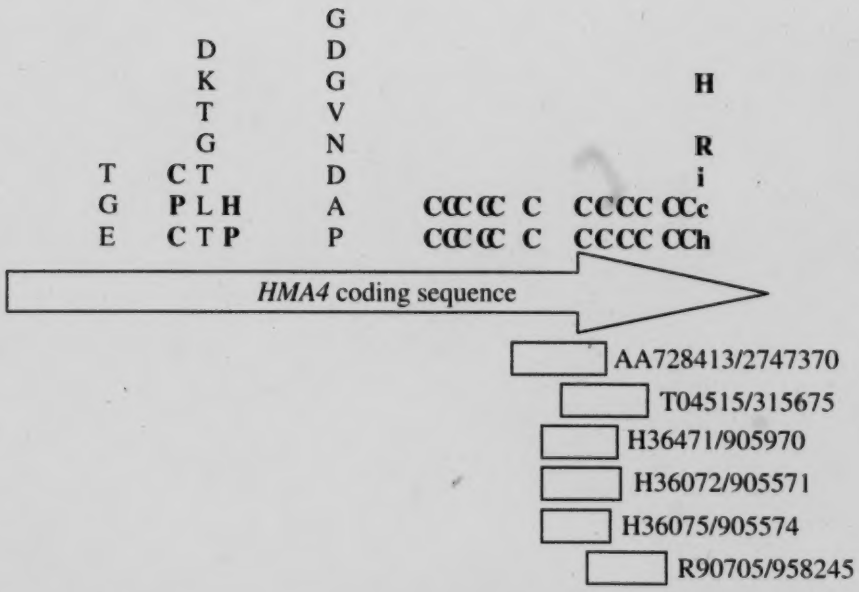


**Figure 2.10 Expressed sequence tags (ESTs) that correspond to the *HMA4* coding sequence.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *HMA4* coding sequences are indicated above the arrow that represents the coding sequence. The ESTs, which were identified by a variety of BLAST searches (Sections 2.2.1.2 and 2.2.1.3.3), are positioned relative to the *HMA4* coding sequence.

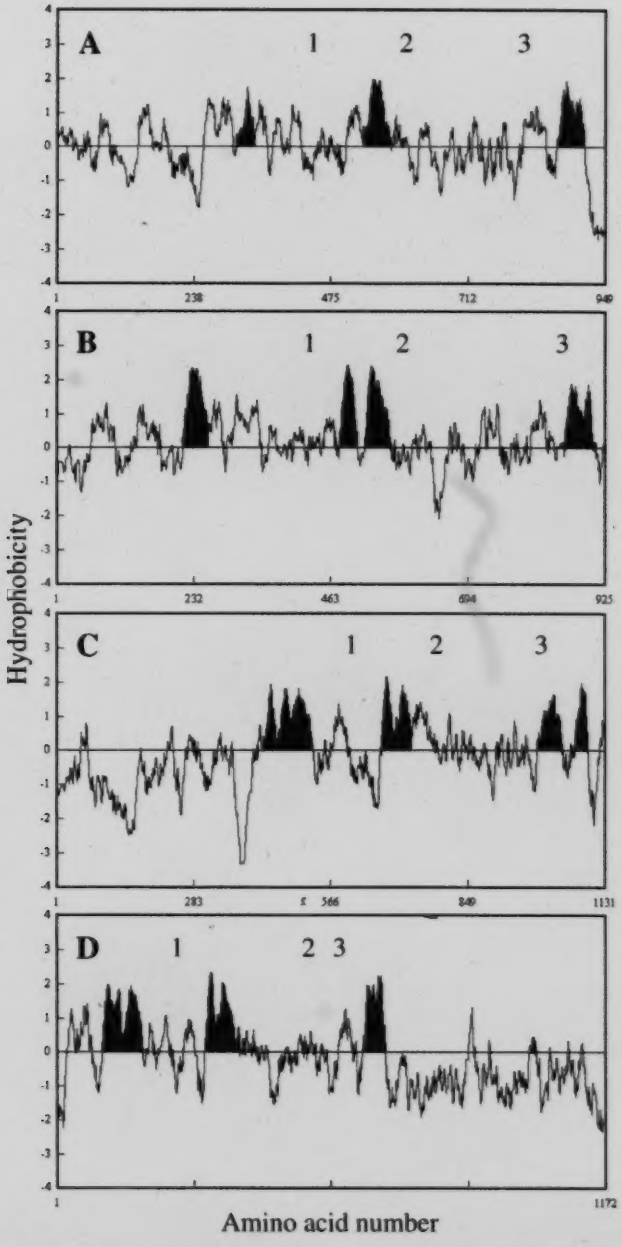


**Figure 2.11 Hydrophobicity profiles for PAA1 (A), RAN1 (B), HMA1 (C), and HMA4 (D).** The hydrophobicity profile option of the DNAMAN program and a window of 19 amino acids was used to create the hydrophobicity profiles. The amino acid segments that likely represent membrane-spanning domains were identified using the criteria of Kyte and Doolittle (1982) and were highlighted with black. The numbers indicate the location of the TGE (1), DKTGT[LIVM][TIS] (2), and GDGxNDxP (3) motifs.

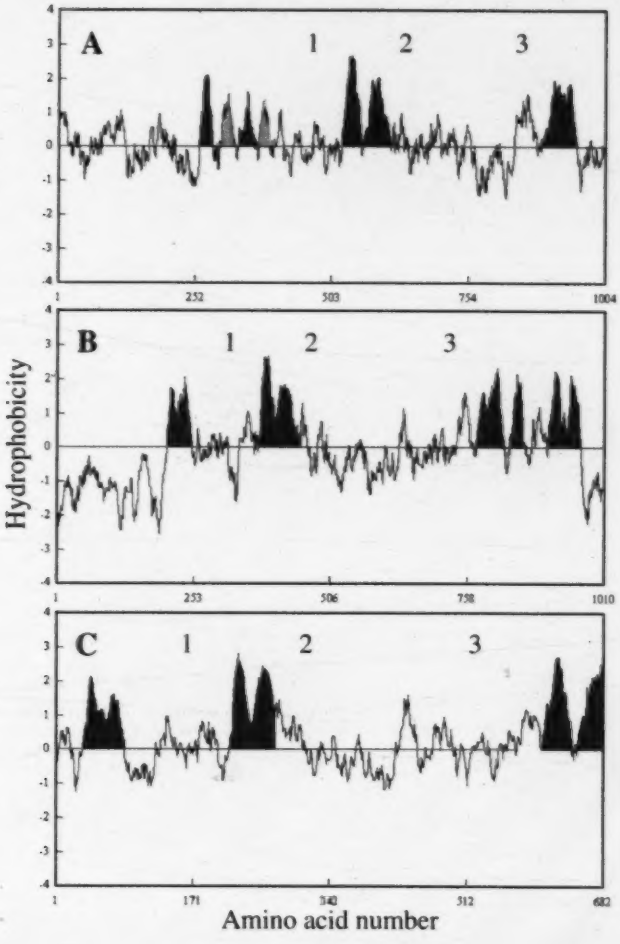


**Figure 2.12 Hydrophobicity profiles for Ccc2p (A; $P_1$-ATPase), Pma2p (B; $P_2$-ATPase), and KdpB (C; $P_3$-ATPase).** The hydrophobicity profile option of the DNAMAN program and a window of 19 amino acids was used to create the hydrophobicity profiles. The amino acid segments that likely represent membrane-spanning domains were identified using the criteria of Kyte and Doolittle (1982) and were highlighted with black. The regions that were identified by Lutsenko and Kaplan (1995), but do not match the criteria of Kyte and Doolittle (1982), were highlighted with grey. The numbers indicate the location of the TGE (1), DKTGT[LIVM][TIS] (2), and GDGxNDxP (3) motifs. The amino acid sequence for Ccc2p, Pma2p, and KdpB are from *S. cerevisiae* (L36317/538514), *S. pombe* (M60471/173430), and *E. coli* (K02670/2772547), respectively.

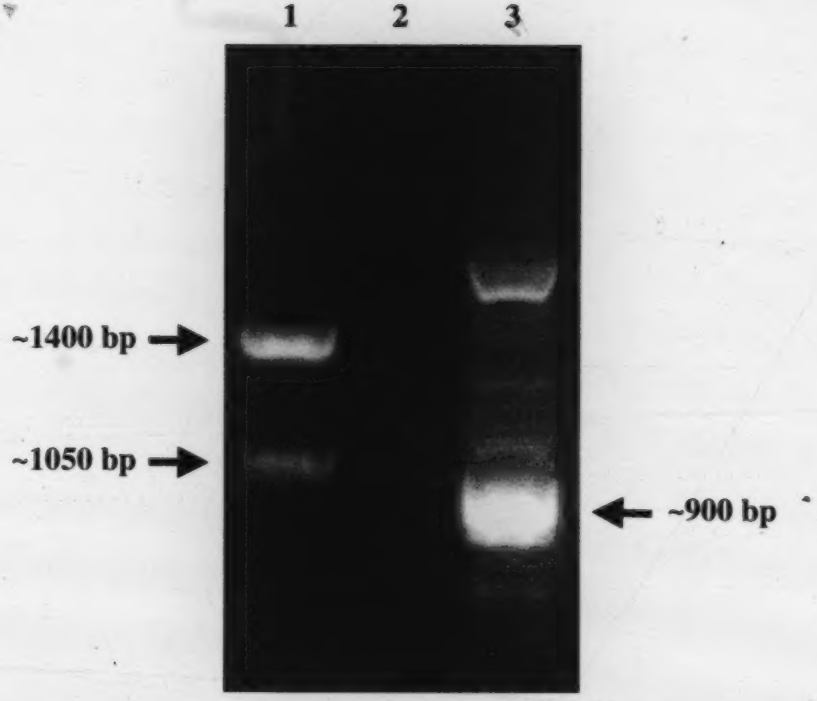


**Figure 2.13 Polymerase chain reaction (PCR) products amplified using the PCR-S/PCR-A primer pair.** Lane 1 contains the products that were amplified when *Arabidopsis* genomic DNA was used as the template. Fragments from the ~1050 and ~1400 bp bands were cloned, partially sequenced, and named PCR-At1 and PCR-At2, respectively. Lane 3 contains products that were amplified when *B. napus* genomic DNA was used as the template. Fragments from the ~900 bp band were cloned and it was subsequently discovered that this band actually contained at least two different PCR fragments, which were ~1000 and ~1250 bp in size. The ~1000 bp product was named PCR-Bn1, while the ~1250 bp product was name PCR-Bn2. Lane 2 was a negative control and did not contain a template.

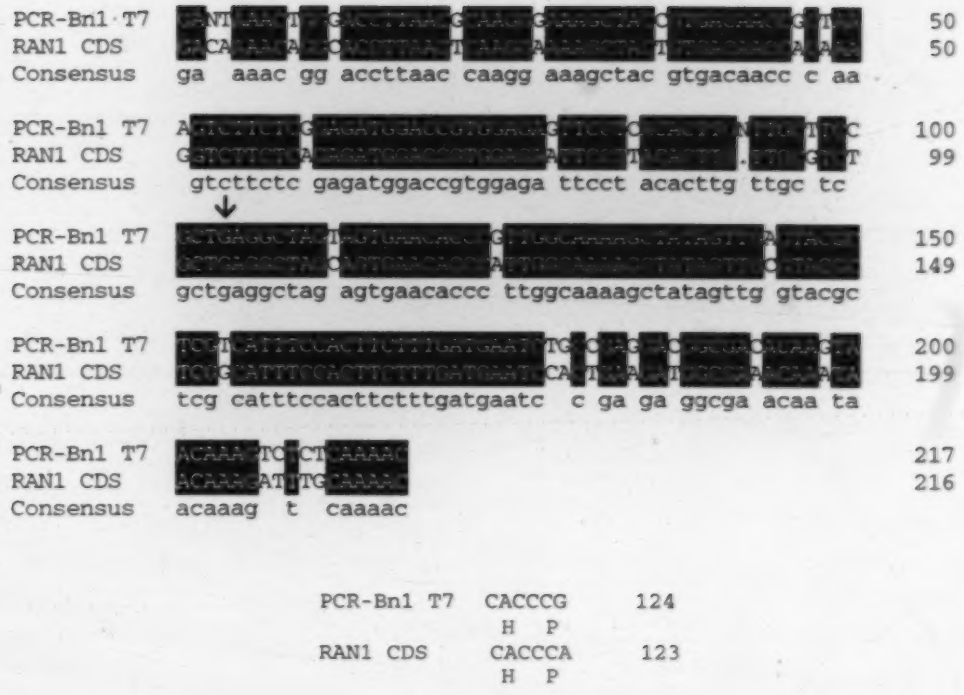


**Figure 2.14 The partial coding sequence (CDS) of PCR-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence.** The PCR-Bn1 product, which was produced using the PCR-S/PCR-A primer pair and *B. napus* genomic DNA, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction contained 83 nucleotides that correspond to a putative intron. In order to compare the sequencing results to the *RAN1* coding sequence, this putative intron sequence was removed. The arrow (↓) indicates where the putative intron was located. The nucleotides that encode the HP locus are presented below the alignment.

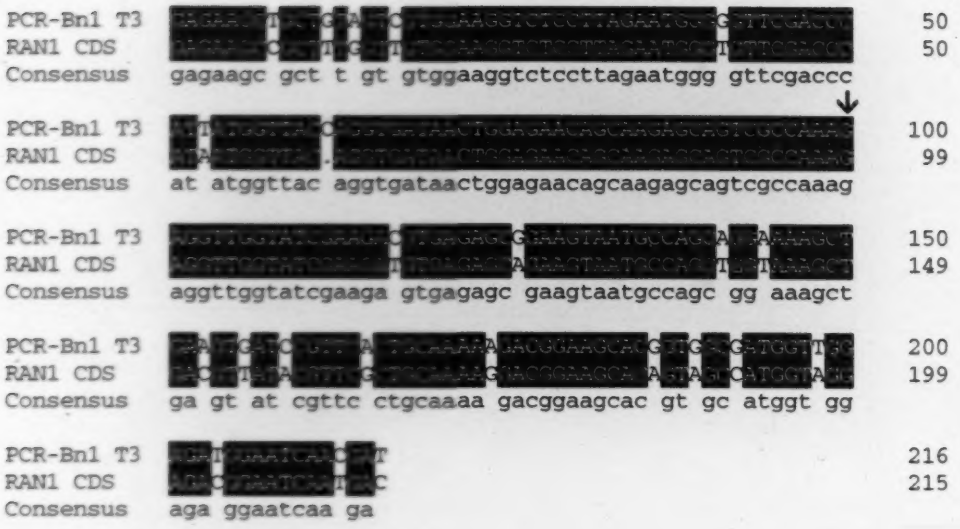


**Figure 2.15 The partial coding sequence (CDS) of PCR-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence.** The PCR-Bn1 product, which was produced using the PCR-S/PCR-A primer pair and *B. napus* genomic DNA, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction contained 84 nucleotides that correspond to a putative intron. In order to compare the sequencing results to the *RAN1* coding sequence (CDS), this putative intron sequence was removed. The arrow (↓) indicates where the putative intron was located.

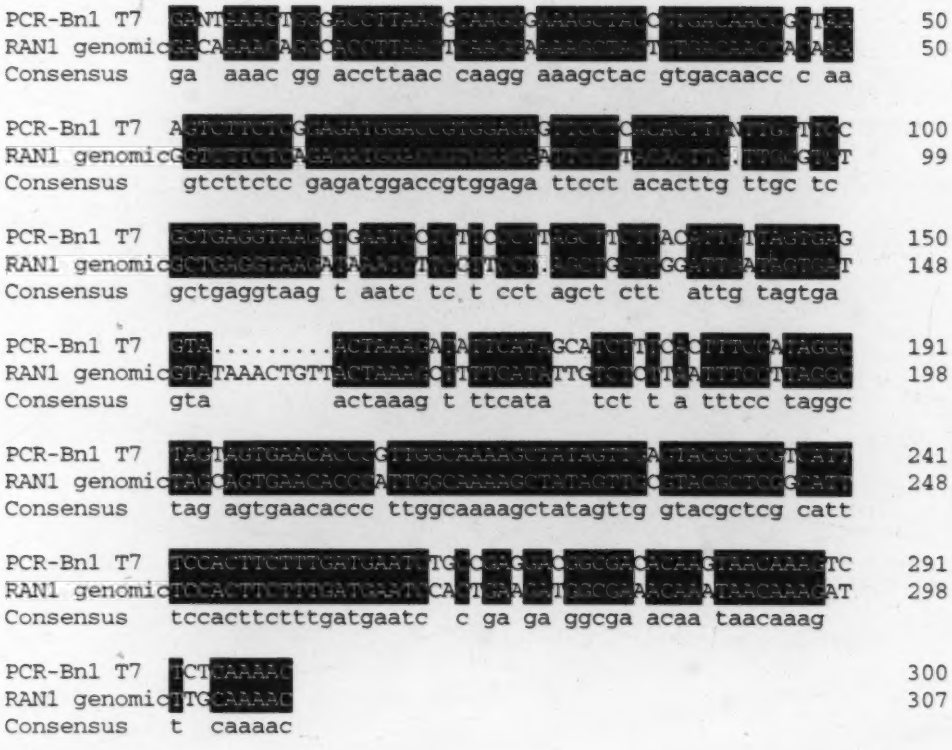


**Figure 2.16 The partial sequence of PCR-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* genomic sequence.** The PCR-Bn1 product, which was produced using the PCR-S/PCR-A primer pair and *B. napus* genomic DNA, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic sequence.

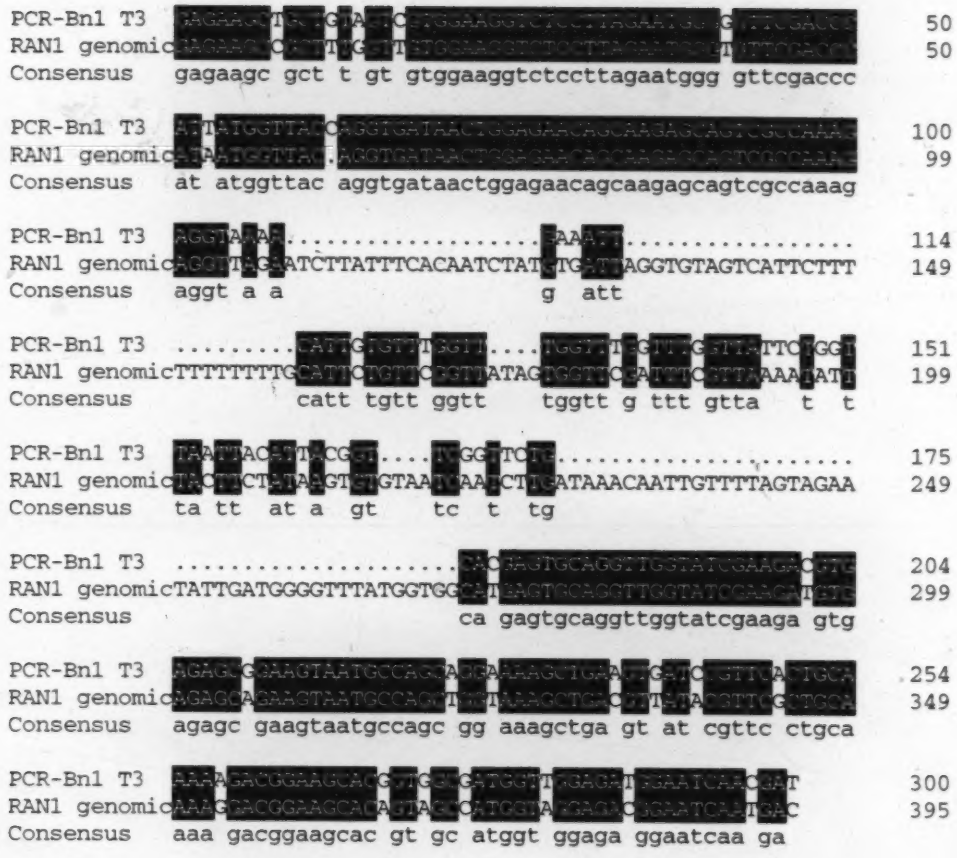
PCR-Bn1 T3
RAN1 genomic
Consensus
gagaagc gct t gt gtggaaggtctccttagaatggg gttcgaccc
at atggttac aggtgataactggagaacagcaagagcagtcgccaaag
aggt a a g att
ca gagtgcaggttggtatcgaaga gtg
agagc gaagtaatgccagc gg aaagctga gt at cgttc ctgca
aaa gacggaagcac gt gc atggt ggaga ggaatcaa ga

**Figure 2.17 The partial sequence of PCR-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* genomic sequence.** The PCR-Bn1 product, which was produced using the PCR-S/PCR-A primer pair and *B. napus* genomic DNA, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic sequence.

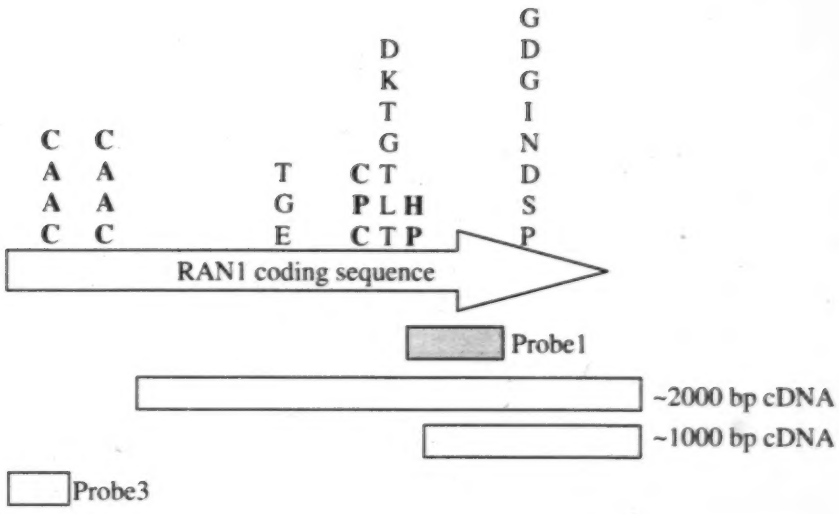


**Figure 2.18 The position of the probes that were produced to screen the amplified *B. napus* library for cDNA(s) that encode a RAN1 homologue and the partial clones that were isolated using Probe1.** The P-type and Type IB (**Bold**) amino acid motifs that are encoded by the *RAN1* coding sequence are indicated above the arrow that represents the coding sequence. Probe1, the ~2000 bp cDNA, the ~1000 bp cDNA, and Probe3 are positioned relative to the *RAN1* coding sequence. The ~2000 bp and ~1000 bp cDNAs were isolate when Probe1 was used to screen the amplified *B. napus* cDNA library. Probe3 was designed to screen the amplified *B. napus* cDNA library for RAN1 homolgues, while avoiding the ~2000 bp and ~1000 bp cDNAs that represent partial clones. Probe1 was derived from a PCR product that was amplified using the Probe1-S/Probe1-A primer pair and the PCR-Bn1 product. The rectangle representing Probe1 has been shaded to indicate that Probe1 contains an intron that is not reflected in its alignment with the *RAN1* coding sequence or the ~2000 and ~1000 bp cDNAs. Probe3 was derived from the Probe3-Bn1 PCR product, which was amplified using the Probe3-S/Probe3-A primer pair and *B. napus* genomic DNA.

**Figure 2.19 The Probe3-Bn1 and Probe3-Bn2 sequences aligned with the appropriate region of the *RAN1* genomic or coding sequence.** The Probe3-Bn1 and Probe3-Bn2 products, which were produced using the Probe3-S/Probe3-A primer pair and *B. napus* genomic DNA, were completely sequenced using the T7 primer. The nucleotides from these sequencing reactions were aligned with the appropriate region of the *RAN1* genomic or coding sequence. There are no introns in the Probe3-Bn1 or Probe3-Bn2 products. Thus, aligning the Probe3-Bn1 and Probe3-Bn2 sequences with the appropriate region of the *RAN1* genomic and coding sequences produces the same results.

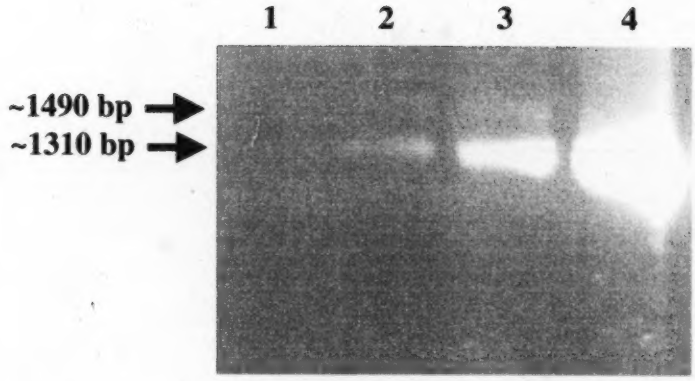


**Figure 2.20 Reverse transcriptase-polymerase chain reaction (RT-PCR) products.** A series of RT-PCRs were setup in which first strand cDNA was produced with the Probe1-A (lane 2), RT3-A (lane 3), or RT1/2-A (lane 4) primers and the final products were amplified using the RT2-S and RT1/2-A primers. Lane 1 was a negative control that did not contain a template. The ~1490 bp band likely resulted from genomic DNA contamination. The RT-Bn1, RT-Bn2, and RT-Bn3 products were cloned from the ~1310 bp bands.

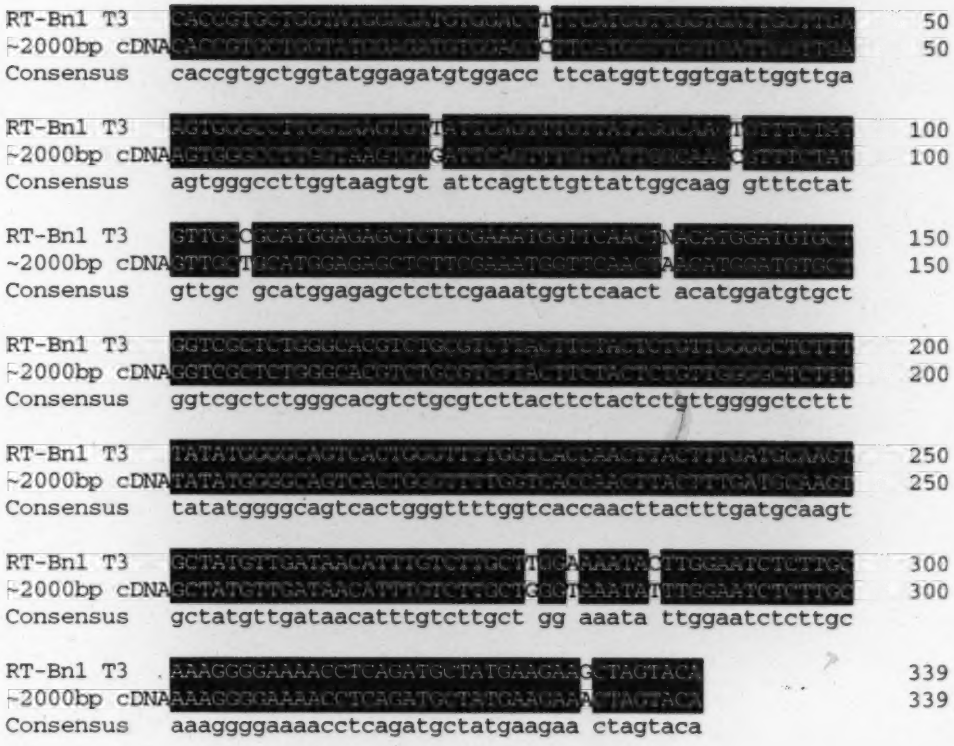


**Figure 2.21 The partial sequence of RT-Bn1 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer.** The RT-Bn1 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the RT3-A primer, was partially sequenced using the T3 primer. The ~2000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T7 primer. The maximum number of overlapping nucleotides for which there was accurate sequencing information were aligned.

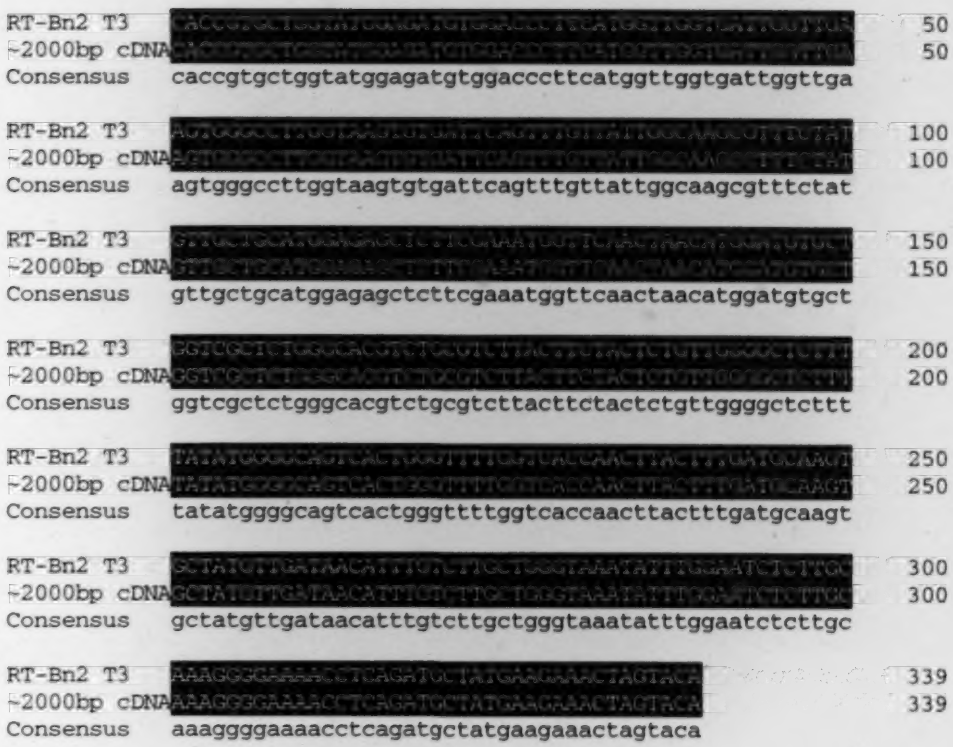


**Figure 2.22 The partial sequence of RT-Bn2 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer.** The RT-Bn2 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the Probe1-A primer, was partially sequenced using the T3 primer. The ~2000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T7 primer. The maximum number of overlapping nucleotides for which there was accurate sequencing information were aligned.

**Figure 2.23 The partial sequence of RT-Bn3 that was produced with the T3 primer aligned with the partial sequence of the ~2000 bp cDNA that was produced with the T7 primer.** The RT-Bn3 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the RT1/2-A primer, was partially sequenced using the T3 primer. The ~2000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T7 primer. The maximum number of overlapping nucleotides for which there was accurate sequencing information were aligned.

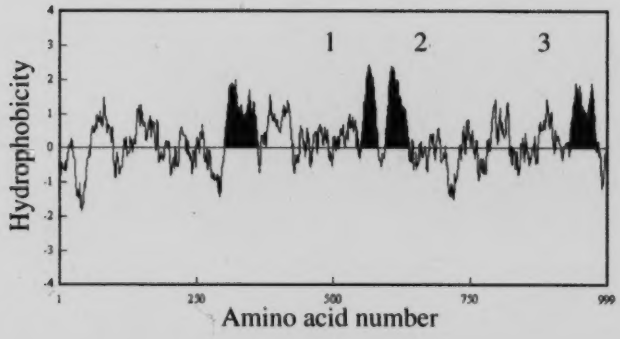


**Figure 2.24 Hydrophobicity profile for the protein encoded by the *B. napus* cDNA.** A Hydrophobicity profile was created for the protein encoded by the *B. napus* cDNA using the hydrophobicity profile option of the DNAMAN program and a window of 19 amino acids. The amino acid segments that likely represent membrane-spanning domains were identified using the criteria of Kyte and Doolittle (1982) and were highlighted with black. The numbers indicate the location of the TGE (1), DKTGT[LIVM][TIS] (2), and GDGxNDxP (3) motifs.

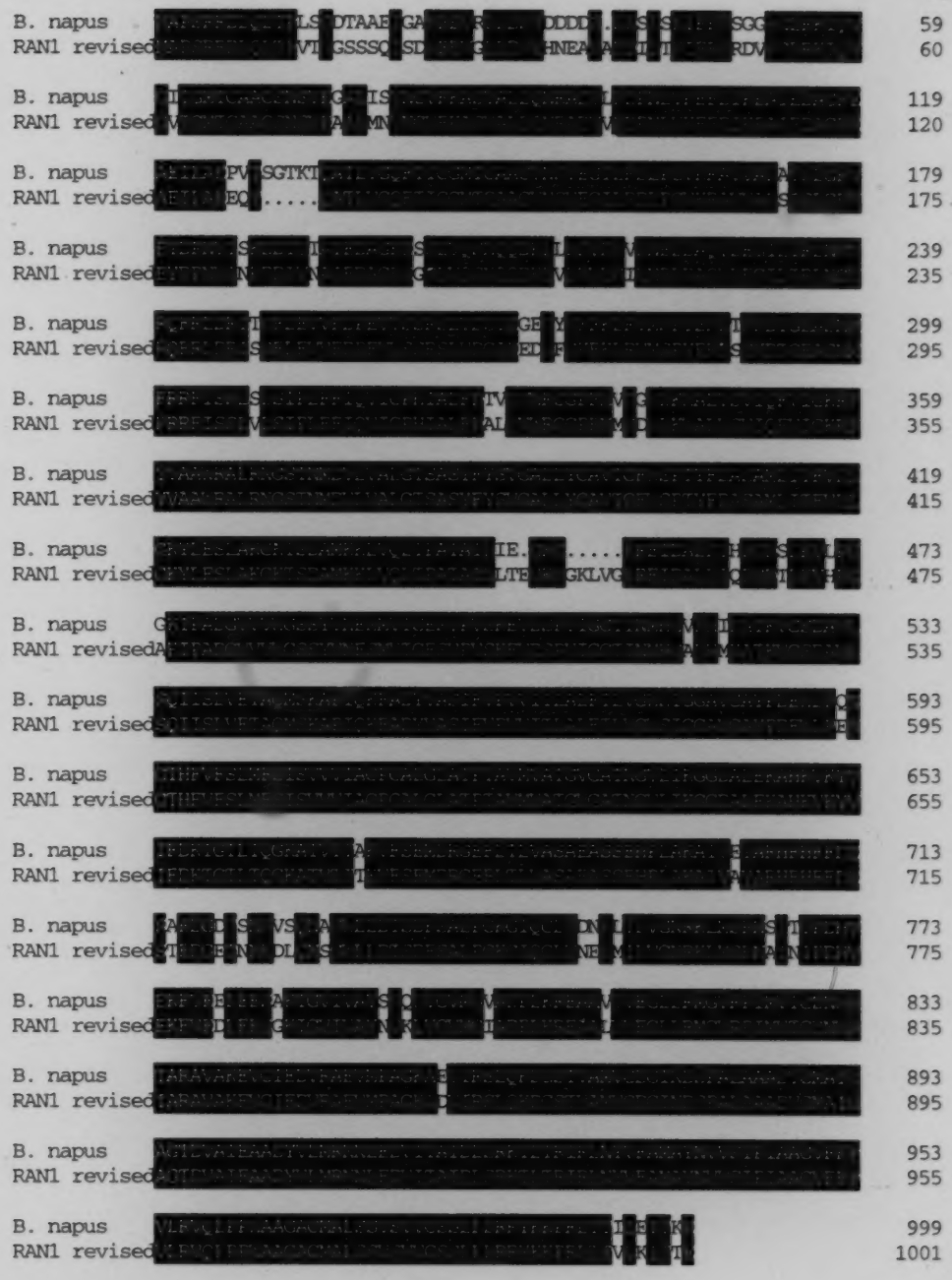


**Figure 2.25 The protein encoded by the *B. napus* cDNA aligned with the updated RAN1 amino acid sequence (AAC79141/6850337).** The amino acid sequence that is encoded by the *B. napus* cDNA was predicted using the translation tool of the DNAMAN program.

**Table 2.1 Polymerase chain reaction (PCR) primer sequences.**

| Primer name | Primer sequence 5' to 3' |
|---|---|
| PCR-S | GACAAAAC(AT)GG(CG)AC(AC)(CT)T |
| PCR-A | (AG)TC(AG)TT(GT)ATTCC(AG)TC(CT)CC |
| Probe1-S | AGTGAACACCCATTGGCAA |
| Probe1-A | CTTTGGCGACTGCTCTTGC |
| Probe2-S | CCTCAAGATTCAAGGTTAG |
| Probe2-A | TAGTCTGACCATTATGAG |
| Probe3-S | AGACGGGATTTACAGCTTACT |
| Probe3-A | CGGCTCGATTCTGTAACAAAG |
| Probe4-S | TCAGCTTTCTCAACCGTCA |
| Probe4-A | TACTCCTGAAAGACGGCG |
| Probe5-S | ATGGAACCTGCAACTCTTACTC |
| Probe5-A | GCACGTAGACGAATCGACCG |
| Probe6-S | TGGATGAGTTTGGCGAGAAT |
| Probe6-A | CTTGCACAGGAACTAATAGC |
| RT1-S | CCGATTTCTCAACCTTCCCG |
| RT1/2-A | ATCGCTGTTGCTGGAGTAAGT |
| RT2-S | AAGAAATGGCGCCGAGTAGAC |
| RT3-A | ATGTATCACCAGGCTGAA |
| Seq1-S | TGCGCCGCTTGTTCTAATTCCG |
| Seq2-S | CCCTGGAGTGAAACGAGCTG |
| Seq3-A | CTCGAACTCACAACCTCAGGATC |
| Seq4-S | GGTTGATGGGATTGAAGAAGATGG |
| Seq5-S | CATGGATGTGCTGGTCGCTCTG |
| Seq6-S | CTGCAGATGGTGTTGTGGTGTG |
| Seq7-A | GAACTCTCCACGGTCCATCTC |
| Seq8-A | CTCCTGTCTTTGCGCTTTCTTC |
| Seq9-A | GTCTCGGCTTCTTGTATCTTCTAA |
| T7 | GTAATACGACTCACTATAGGGC |
| T3 | AATTAACCCTCACTAAAGGG |

**Table 2.2 Polymerase chain reaction products that could be produced by primers based on the DKTGT[LIVM][TIS] and GDGxNDxP amino acid sequences.**
*Arabidopsis* sequences from the databases were used to predict the potential genomic and coding sequence products that might be produced using primers that were based on the DKTGT[LIVM][TIS] and GDGxNDxP amino acid sequences. The *PAA1*, *RAN1*, *HMA1*, and *HMA4* sequences were used to predict products that would correspond to putative $P_{1B}$-ATPases. Potential products from sequence that belonged to P-type ATPase subfamilies other than the Type IB subfamily were predicted using the full-length *Arabidopsis* sequences *ACA3*/U93845/1943750, *PEA1*/L08468/509809, *AHA2*/SwissProt P19456, *AHA3*/SwissProt P20431, *AHA9*/X73676/471289, *AHA10*/ S74033/765353 that had been identified by Axelsen and Palmgren (1998).

| Template name | P-type ATPase subfamily | Size of the product predicted from the coding sequence (bp) | Size of the product predicted from the genomic sequence (bp) | Number of introns in the genomic product | Average intron size in the genomic product (bp) |
|---|---|---|---|---|---|
| *PAA1* | IB | 645 | Unavailable | Unavailable | Unavailable |
| *RAN1* | IB | 672 | 1038 | 4 | 92 |
| *HMA1* | IB | 702 | 1503 | 4 | 200 |
| *HMA4* | IB | 615 | 867 | 3 | 84 |
| *ACA3* | IIA | 1059 | Unavailable | Unavailable | Unavailable |
| *PEA1* | IIB | 921 | Unavailable | Unavailable | Unavailable |
| *AHA2* | IIIA | 792 | 1248 | 5 | 91 |
| *AHA3* | IIIA | 792 | 1244 | 5 | 90 |
| *AHA9* | IIIA | 792 | 1039 | 3 | 82 |
| *AHA10* | IIIA | 789 | 1438 | 6 | 108 |

**Table 2.3 Multiple alignment results.** Nucleic acid fragments of interest were either partially or completely sequenced. The asterisk (*) identifies the fragments that were completely sequenced. The sequencing reactions usually provided at least 300 bp of clear insert sequence from each primer. To promote consistency when an insert is only partially sequenced, the sequencing results that were analyzed were limited to the first 300 bp of insert sequence provided by each primer. The sequencing results were analyzed by BLAST searches followed by a comparison to the appropriate region of the top match using the multiple alignment tool of the DNAMAN program. If a putative intron was removed prior to an alignment "CDS" (coding sequence) was used to describe the product. The alignments that correspond to the information in this table can be found in Figures 2.16 to 2.19 in this Chapter and Figures 5.1 to 5.20 in the Appendix.

| Product that was sequenced | Primer used in the sequencing reaction | Sequencing results were aligned with the appropriate region of: | Percent identity (%) | Figure that corresponds to this alignment |
|---|---|---|---|---|
| PCR-At1 | T3 | *RAN1* genomic | 99 | Figure 5.1 |
| PCR-At1 | T7 | *RAN1* genomic | 98 | Figure 5.2 |
| PCR-At2 CDS | T7 | *PAA1* mRNA/CDS | 97 | Figure 5.3 |
| PCR-At2 CDS | T3 | *PAA1* mRNA/CDS | 98 | Figure 5.4 |
| PCR-Bn1 | T7 | *RAN1* genomic | 81 | Figure 2.18 |
| PCR-Bn1 CDS | T7 | *RAN1* CDS | 86 | Figure 2.16 |
| PCR-Bn1 | T3 | *RAN1* genomic | 64 | Figure 2.19 |
| PCR-Bn1 CDS | T3 | *RAN1* CDS | 89 | Figure 2.17 |
| PCR-Bn2 CDS | T3 | *PAA1* mRNA/CDS | 87 | Figure 5.5 |
| PCR-Bn2 CDS | T7 | *PAA1* mRNA/CDS | 86 | Figure 5.6 |
| ~2000 bp cDNA | T7 | *RAN1* CDS | 91 | Figure 5.7 |
| ~2000 bp cDNA | T3 | *RAN1* CDS | 92 | Figure 5.8 |
| ~1000 bp cDNA | T7 | *RAN1* CDS | 90 | Figure 5.9 |
| Probe3-Bn1* | T7 | *RAN1* genomic or CDS | 77 | Figure 5.10 |
| Probe3-Bn2* | T7 | *RAN1* genomic or CDS | 82 | Figure 5.11 |
| Probe4-Bn1* | T3 | *PAA1* genomic or CDS | 68 | Figure 5.12 |
| Probe5-At1* | T7 | *HMA1* genomic or CDS | 97 | Figure 5.13 |
| Probe6-At1* | T3 and T7 | *HMA1* genomic | 93 | Figure 5.14 |
| Probe6-Bn1* | T3 and T7 | *HMA1* genomic | 78 | Figure 5.15 |
| Probe6-Bn1 CDS* | T3 and T7 | *HMA1* CDS | 93 | Figure 5.16 |
| RT-Bn1 | T7 | *RAN1* CDS | 79 | Figure 5.17 |
| RT-Bn1 | T3 | *RAN1* CDS | 92 | Figure 5.18 |
| RT-Bn2 | T7 | *RAN1* CDS | 82 | Figure 5.19 |
| RT-Bn2 | T3 | *RAN1* CDS | 93 | Figure 5.20 |

## 2.5 Literature cited

**Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. J Mol Biol **215(3)**:403-410

**Axelsen KB, Palmgren MG** (1998) Evolution of substrate specificities in the P-type ATPase superfamily. J Mol Evol **46(1)**:84-101

**Axelsen KB, Palmgren MG** (2001) Inventory of the superfamily of P-type ion pumps in *Arabidopsis*. Plant Physiol **126(2)**:696-706

**Bairoch A, Apweiler R** (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res **28(1)**:45-48

**Barker WC, Garavelli JS, Hou Z, Huang H, Ledley RS, McGarvey PB, Mewes HW, Orcutt BC, Pfeiffer F, Tsugita A, Vinayaka CR, Xiao C, Yeh LS, Wu C** (2001) Protein Information Resource: a community resource for expert annotation of protein data. Nucleic Acids Res **29(1)**:29-32

**Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL** (2000) GenBank. Nucleic Acids Res **28(1)**:15-18

**Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE** (2000) The Protein Data Bank. Nucleic Acids Res 2000 **28(1)**:235-242

**Boguski MS, Lowe TM, Tolstoshev CM** (1993) dbEST--database for "expressed sequence tags". Nat Genet **4(4)**:332-333

**Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW** (1993) The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. Nat Genet **5(4)**:327-337

**Chelly J, Tumer Z, Tonnesen T, Petterson A, Ishikawa-Brush Y, Tommerup N, Horn N, Monaco AP** (1993) Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. Nat Genet **3(1)**:14-19

**Forbes JR** (2000) Functional analysis of the copper-transporting P-type ATPase, ATP7B, defective in Wilson disease. PhD thesis. University of Alberta, Canada, 2000 159 pp

**Forbes JR, Cox DW** (1998) Functional characterization of missense mutations in *ATP7B*: Wilson disease mutation or normal variant? Am J Hum Genet **63(6)**:1663-1674

**Gitschier J, Moffat B, Reilly D, Wood WI, Fairbrother WJ** (1998) Solution structure of the fourth metal-binding domain from the Menkes copper-transporting ATPase. Nat Struct Biol **5(1)**:47-54

**Hirayama T, Kieber JJ, Hirayama N, Kogan M, Guzman P, Nourizadeh S, Alonso JM, Dailey WP, Dancis A, Ecker JR** (1999) RESPONSIVE-TO-ANTAGONIST1, a Menkes/Wilson disease-related copper transporter, is required for ethylene signaling in *Arabidopsis*. Cell **97(3)**:383-393

**Huffman DL, O'Halloran TV** (2001) Function, structure, and mechanism of intracellular copper trafficking proteins. Annu Rev Biochem **70**:677-701

**Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD, Gitlin JD** (1997) Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. J Biol Chem **272(34)**:21461-21466

**Iida M, Terada K, Sambongi Y, Wakabayashi T, Miura N, Koyama K, Futai M, Sugiyama T** (1998) Analysis of functional domains of Wilson disease protein (ATP7B) in *Saccharomyces cerevisiae*. FEBS Lett **428(3)**:281-285

**Kyte J, Doolittle RF** (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157(1)**:105-132

**Lutsenko S, Kaplan JH** (1995) Organization of P-type ATPases: significance of structural diversity. Biochemistry **34(48)**:15607-15613

**Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M** (1998) *Arabidopsis thaliana*: a model plant for genome analysis. Science **282(5389)**:679-682

**Mercer JF, Livingston J, Hall B, Paynter JA, Begy C, Chandrasekharappa S, Lockhart P, Grimes A, Bhave M, Siemieniak D, *et al.*** (1993) Isolation of a partial candidate gene for Menkes disease by positional cloning. Nat Genet **3(1)**:20-25

**Mewes HW, Frishman D, Gruber C, Geier B, Haase D, Kaps A, Lemcke K, Mannhaupt G, Pfeiffer F, Schuller C, Stocker S, Weil B** (2000) MIPS: a database for genomes and protein sequences. Nucleic Acids Res **28(1)**:37-40

**Nucifora G, Chu L, Misra TK, Silver S** (1989) Cadmium resistance from *Staphylococcus aureus* plasmid pI258 *cadA* gene results from a cadmium-efflux ATPase. Proc Natl Acad Sci U S A **86(10)**:3544-3548

**Palmgren MG, Harper JF** (1999) Pumping with plant P-type ATPases. J Exp Bot **50**:883-893

**Payne AS, Gitlin JD** (1998) Functional expression of the Menkes disease protein reveals common biochemical mechanisms among the copper-transporting P-type ATPases. J Biol Chem **273(6)**:3765-3770

**Sambongi Y, Wakabayashi T, Yoshimizu T, Omote H, Oka T, Futai M** (1997) *Caenorhabditis elegans* cDNA for a Menkes/Wilson disease gene homologue and its function in a yeast *CCC2* gene deletion mutant. J Biochem (Tokyo) **121(6)**:1169-1175

**Sambrook J, Fritsch EF, Maniatis T** (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory, 1989.

**Schuler GD, Epstein JA, Ohkawa H, Kans JA** (1996) Entrez: molecular biology database and retrieval system. Methods Enzymol **266**:141-162

**Silver S, Nucifora G, Chu L, Misra TK** (1989) Bacterial resistance ATPases: primary pumps for exporting toxic cations and anions. Trends Biochem Sci **14(2)**:76-80

**Silver S, Nucifora G, Phung LT** (1993) Human Menkes X-chromosome disease and the *Staphylococcal* cadmium-resistance ATPase: a remarkable similarity in protein sequences. Mol Microbiol **10(1)**:7-12

**Solioz M, Odermatt A** (1995) Copper and silver transport by CopB-ATPase in membrane vesicles of *Enterococcus hirae*. J Biol Chem **270(16)**:9217-9221

**Solioz M, Vulpe C** (1996) CPx-type ATPases: a class of P-type ATPases that pump heavy metals. Trends Biochem Sci **21(7)**:237-241

**Stoesser G, Baker W, van den Broek A, Camon E, Garcia-Pastor M, Kanz C, Kulikova T, Lombard V, Lopez R, Parkinson H, Redaschi N, Sterk P, Stoehr P, Tuli MA** (2001) The EMBL nucleotide sequence database. Nucleic Acids Res **29(1)**:17-21

**Tabata K, Kashiwagi S, Mori H, Ueguchi C, Mizuno T** (1997) Cloning of a cDNA encoding a putative metal-transporting P-type ATPase from *Arabidopsis thaliana*. Biochim Biophys Acta **1326(1)**:1-6

**Tanzi RE, Petrukhin K, Chernov I, Pellequer JL, Wasco W, Ross B, Romano DM, Parano E, Pavone L, Brzustowicz LM, *et al.*** (1993) The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. Nat Genet **5(4)**:344-350

**Tateno Y, Miyazaki S, Ota M, Sugawara H, Gojobori T** (2000) DNA data bank of Japan (DDBJ) in collaboration with mass sequencing teams. Nucleic Acids Res **28(1)**:24-26

**Tsai KJ, Yoon KP, Lynn AR** (1992) ATP-dependent cadmium transport by the *cadA* cadmium resistance determinant in everted membrane vesicles of *Bacillus subtilis*. J Bacteriol **174(1)**:116-121

**Voskoboinik I, Brooks H, Smith S, Shen P, Camakaris J** (1998) ATP-dependent copper transport by the Menkes protein in membrane vesicles isolated from cultured Chinese hamster ovary cells. FEBS Lett **435(2-3)**:178-182

**Vulpe C, Levinson B, Whitney S, Packman S, Gitschier J** (1993) Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nat Genet **3(1)**:7-13

**Wheeler DL, Church DM, Lash AE, Leipe DD, Madden TL, Pontius JU, Schuler GD, Schriml LM, Tatusova TA, Wagner L, Rapp BA** (2001) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res **29(1)**:11-16

**Williams LE, Pittman JK, Hall JL** (2000) Emerging mechanisms for heavy metal transport in plants. Biochim Biophys Acta **1465(1-2)**:104-126

**Woeste KE, Kieber JJ (2000)** A strong loss-of-function mutation in *RAN1* results in constitutive activation of the ethylene response pathway as well as a rosette-lethal phenotype. Plant Cell **12(3)**:443-455

# 3 Functional complementation of the *Saccharomyces cerevisiae ccc2* mutant by a *Brassica napus* cDNA (AY045772/15636780) encoding a putative copper-transporting P-type ATPase

## 3.1 Introduction

The budding yeast *Saccharomyces cerevisiae* is a powerful eukaryotic model system. Some of the benefits of working with *S. cerevisiae* include rapid growth, dispersed cells, and a completely sequenced genome. The development of a wealth of genetic tools (Sherman, 1997) has helped to strengthen the role of *S. cerevisiae* as a model organism and as an expression system for the functional analyzes of genes from other species.

*Saccharomyces cerevisiae* strains that carry disruptions in the *CCC2* gene have become an important tool for studying putative $Cu^{2+}$-ATPases from other organisms. The ability of a putative $Cu^{2+}$-ATPase to rescue the *ccc2* mutant has been interpreted as evidence suggesting that the protein transports copper (Hung *et al.*, 1997; Sambongi *et al.*, 1997; Forbes and Cox, 1998; Iida *et al.*, 1998; Payne and Gitlin, 1998; Hirayama *et al.*, 1999; Forbes, 2000). The *Brassica napus* cDNA cloned in this study (AY045772/15636780) has been predicted to encode a putative $Cu^{2+}$-ATPase. Thus, a *ccc2* complementation assay will be used to test this prediction.

Complementation of *ccc2* mutants has been demonstrated in a number of ways. As discussed previously, disruption of the *CCC2* gene leads to a loss of Fet3p activity and high-affinity iron uptake (Section 1.5). Restoration of high-affinity iron uptake can be demonstrated in several ways. High-affinity iron uptake has been monitored directly using radioactive iron (Hirayama *et al.*, 1999). In addition, cells that are deficient in high-affinity iron uptake cannot grow under iron-limited conditions; thus, growth on iron-

limited medium has been used to demonstrate complementation (Forbes and Cox, 1998; Iida *et al.*, 1998; Forbes, 2000). A third possibility arises from the fact that copper and iron are required for respiration as cofactors in cytochrome-*c* oxidase. A deficiency in high-affinity iron uptake, therefore, results in a respiration-deficiency phenotype (Yuan *et al.*, 1995). Respiration-deficiency can be demonstrated by an inability to grow on non-fermentable, glycerol-based medium. Studies have, consequently, utilized glycerol-based medium to study complementation of *ccc2* mutants (Yuan *et al.*, 1995; Sambongi *et al.*, 1997). The Ccc2 protein (or a functional homolog) is required to deliver copper to Fet3p. Only copper-containing Fet3p is active and an oxidase assay that demonstrates Fet3p activity provides a further method of demonstrating complementation of *ccc2* mutants (Hung *et al.*, 1997; Forbes and Cox, 1998; Payne and Gitlin, 1998; Hirayama *et al.*, 1999; Forbes, 2000). Finally, *in vivo* copper incorporation into Fet3p has been demonstrated by pulse-labeling with $^{64}$Cu followed by autoradiography of membrane extracts that have been separated by nonreducing SDS-PAGE (Hung *et al.*, 1997; Payne and Gitlin, 1998).

A complementation assay that utilizes iron-limited medium will be used in this study. This assay provides a number of advantages over other methods of assaying for *ccc2* complementation. In contrast to the high-affinity iron uptake analysis and pulse-labeling with $^{64}$Cu, the iron-limited medium assay does not involve radioactive isotopes. The iron-limited medium provides a more complete growth inhibition of *ccc2* mutants than the glycerol-based medium (Forbes, 2000). It is, therefore, easier to observe complementation on iron-limited medium than on glycerol-based medium. The iron-limited medium assay also provides a relatively fast and easy method of analyzing complementation of *ccc2* mutants. Growth data in iron-limited medium has also been found to agree well with the results of the more complicated oxidase assay (Forbes and Cox, 1998). Finally, a version of the iron-limited medium assay was developed at the University of Alberta (Forbes and Cox, 1998; Forbes, 2000) as one component of an indirect method for testing the degree to which Wilson disease mutations in the *ATP7B* cDNA affect protein function. Thus, guidance was close by. The final objective of this study was, therefore, to use this complementation assay to determine if the *B. napus* cDNA does in fact encode a protein that is able to transport copper.

## 3.2 Materials and methods

### 3.2.1 Construction of the yeast expression vector

The *B. napus* cDNA encoding the putative $Cu^{2+}$-ATPase was sub-cloned into the pYES3 vector (Invitrogen) using standard molecular biology techniques (Section 2.2.2.1). The *B. napus* cDNA was liberated from the pBluescript SK- vector using the *HindIII* and *BamHI* restriction enzymes (Figure 3.1). The appropriate restriction fragment was cloned into a *HindIII* / *BamHI*-digested pYES3 vector and the resulting vector was named "pYES3-cDNA" (Figure 3.2). Prior to sub-cloning the *B. napus* cDNA into the pYES3 vector, attempts had been made to sub-clone the *B. napus* cDNA into the pG3 vector (Schena *et al.*, 1991); unfortunately, these attempts were unsuccessful. The pG3 vector was used by Forbes and Cox (1998) and had been obtained from the laboratory of Dr. Diane Cox at the University of Alberta (Edmonton, Alberta, Canada). Another pG3 vector into which the *ATP7B* cDNA had been sub-cloned (Forbes and Cox, 1998) was also obtained from the laboratory of Dr. Diane Cox and will be referred to as "pG3-ATP7B". All manipulations and amplification of these vectors were performed in *Escherichia coli*.

The pYES3 and pG3 vectors are similar in a number of important ways. Both are shuttle vectors that can be used to transform *E. coli* and *S. cerevisiae*. The *TRP1* gene is found on both vectors and enables the isolation of *S. cerevisiae* cells containing the vector. An ampicillin-resistance gene ($Amp^r$) is also found on both vectors and is involved in the selection of *E. coli* cells that have been transformed with either of the vectors. High-copy numbers of the pG3 and pYES3 vectors are maintained in *S. cerevisiae* cells by the 2μ sequence. Prokaryotic origins of replication mediate the maintenance of high-copy numbers of the vectors in *E. coli* cells. These origins of replication are, however, different in each vector. The pG3 vector contains the pMB1 origin and the pYES3 vector includes the pUC origin. Since the pYES3 and pG3 vectors are yeast expression vectors, they contain components that play an important role in the translation and transcription of the insert. Both vectors contain sequences to stabilize the

mRNA produced from the insert. In the pYES3 and pG3 vectors these sequences are derived from the *CYC1* and *PGK* terminator and polyadenylation sequences, respectively.

Although the pYES3 and pG3 vectors both contain promoters from yeast genes, the promoters are different in each vector, thus the conditions required for transcription of the insert vary between the two vectors. The pYES3 vector contains the *GAL1* promoter and the pG3 vector contains the *GDP* promoter. Transcription from the *GAL1* promoter is induced in the presence of galactose, repressed in glucose-containing media, and neither repressed nor induced by raffinose. In contrast, transcription from the *GDP* promoter is constitutive and does not require induction. Another important difference between the pYES3 and pG3 vectors is the fact that the pG3 vector can be easily converted to the integrative vector pG4 by removing the 2μ sequence. The 2μ sequence is removed from the pG3 vector using two *EcoRI* restriction sites that flank it. The pYES3 vector does not offer a simple method of removing the 2μ sequence. Integrative vectors cannot replicate autonomously and must integrate into the genome by homologous recombination. The integrative vector, therefore, allows single-copy expression of an insert.

Forbes and Cox (1998) demonstrated rescue of the *ccc2* mutant by both single (pG4-ATP7B) and multicopy (pG3-ATP7B) expression of ATP7B. Since Forbes and Cox (1998) developed the complementation assay as an indirect method of testing the degree to which Wilson disease mutations in the *ATP7B* gene affect protein function, they required a quantitative assay and the mutated versions of the *ATP7B* cDNA were integrated as single copies. The purpose of the current study was to determine if the *B. napus* cDNA encodes a protein capable of transporting copper. It was not the intention of this study to quantitatively compare the activity of the *B. napus* protein to any other proteins. Single copy integration of the *B. napus* cDNA was, therefore, not considered to be essential for this study. A number of groups have used multicopy vectors to demonstrate complementation of the *ccc2* mutant (Hung *et al.*, 1997; Sambongi *et al.*, 1997; Forbes and Cox, 1998; Iida *et al.*, 1998; Hiryama *et al.*, 1999). In some cases

comparisons have even been made between the activity of wild-type and mutant alleles that were expressed from multicopy vectors (Hung *et al.*, 1998; Hiryama *et al.*, 1999).

### 3.2.2 General growth conditions and techniques

Well-established techniques (Sherman, 1991) were used in the general growth, maintenance, and manipulation of the *S. cerevisiae* strains used in this study. The term "conventional" will be used to describe all media that was made with commonly accepted constituents (Sherman, 1991). The carbon source in conventional media is glucose and the YNB used in conventional synthetic medium normally contains 0.74 μM ferric chloride and 0.16 μM copper sulfate. Unless otherwise stated, the amino acid supplementation pattern that follows was used to make all conventional or modified synthetic media. All of the amino acids were added when culturing the BJ2168 non-transformed strain. Uracil was the only amino acid omitted when the *ccc2*, *fet3*, and *ctr1* non-transformed strains were grown. The media used for all transformed strains was made with all of the amino acids except tryptophan. When a medium is a modified version of a conventional medium, only the constituents that vary from the conventional medium will be described. All plates were made by adding 2% Bacto-agar (Difco) to the appropriate liquid medium.

### 3.2.3 *Saccharomyces cerevisiae* strains

All *S. cerevisiae* strains were obtained from the laboratory of Dr. Diane Cox at the University of Alberta (Edmonton, Alberta, Canada). The strains were received on the appropriate conventional synthetic plates and had been streaked from glycerol stocks. Glycerol stocks were made for maintenance of the strains in our laboratory. The parental strain is the protease-deficient BJ2168 strain (*MATa pep4-3 prc1-407 prb1-1122 ura3-52 trp1 leu2*; Zubenko *et al.*, 1980) that the Cox laboratory had obtained from Dr. Morrie Manolson at the University of Toronto (Toronto, Ontario, Canada). The *CCC2*, *FET3*, and *CTR1* genes of the BJ2168 strain were independently disrupted to create the *ccc2*, *fet3*, and *ctr1* stains, respectively (Forbes, 2000).

### 3.2.4 *Saccharomyces cerevisiae* transformation

The *S. cerevisiae* strains BJ2168, *ccc2*, *fet3*, and *ctr1* were independently transformed with the pYES3, pYES3-cDNA, pG3, and pG3-ATP7B vectors. A modified lithium acetate method (Elble, 1992) was used for all transformations. The transformants were selected and restreaked on conventional synthetic plates that were supplemented with all of the amino acids except tryptophan. Isolated colonies from the restreaked plates were used to make glycerol stocks. Names were assigned to the transformants that described both the strain that was transformed and the vector that was used in the transformation. For example, when the BJ2168 strain was transformed with the pYES3-cDNA vector it was named "BJ2168/pYES3-cDNA".

### 3.2.5 Iron-limited, iron-sufficient, copper-sufficient, and induction media

A modified version of the iron-limited medium assay developed by Forbes and Cox (1998) was selected for use in this study. The iron-limited medium used by Forbes and Cox (1998) was made with a Yeast Nitrogen Base (YNB) that lacked both copper sulfate and ferric chloride (Bio-101). The iron in the iron-limited medium was then adjusted to 50 µM with ferrous ammonium sulfate and its availability was limited by the addition of 1 mM ferrozine (an iron-specific chelator). The iron-limited medium was also adjusted to contain a final concentration of 1 µM copper sulfate. This level of copper was selected (Forbes, 2000) because it has been shown to be insufficient to reconstitute apoFet3p in the plasma membrane (Yuan *et al.*, 1997). The Fet3 protein can be reconstituted at lower copper concentrations if the pH of the medium is too acidic (Davis-Kaplan *et al.*, 1998). The iron-limited medium was, therefore, supplemented with 50 mM MES buffer (pH 6.1; Forbes, 2000). It should be noted that this was not a copper-limited medium, as the medium did not contain a copper chelator and 1 µM copper sulfate has been shown to be sufficient to support limited growth of a *ctr1* mutant (Dancis *et al.*, 1994). The iron-limited medium was supplemented to 350 µM ferrous

ammonium sulfate to create an iron-sufficient medium. A copper-sufficient medium was produced by supplementing the iron-limited medium to 500 μM copper sulfate. The copper- and iron-supplemented media were used to demonstrate that all of the *ccc2* strains studied by Forbes and Cox (1998) were viable.

The induction, iron-limited, iron-sufficient, and copper-sufficient media used in this study were also modified versions of conventional synthetic medium. The induction medium contained 2% galactose and 1% raffinose in place of the 2% glucose that is normally used in conventional synthetic medium. The iron-limited, iron-sufficient, and copper-sufficient media were based on the media described by Forbes and Cox (1998). In this study, however, the carbon source in the iron-limited, iron-sufficient, and copper-sufficient media was changed from 2% glucose to 2% galactose and 1% raffinose. This carbon source change was required to allow transcription from the pYES3 vector. The iron-limited medium was made with a YNB that did not contain copper and iron (Bio-101), but was supplemented with 50 mM MES buffer (pH 6.1), 1 mM ferrozine, 50 μM ferrous ammonium sulfate, and 1 μM copper sulfate. The copper-sufficient media was made by altering the iron-limited media so that it contained a final concentration of 500 μM copper sulfate. The iron-sufficient medium was made by omitting the ferrozine from the iron-limited media and adjusting ferrous ammonium sulfate to a final concentration of 350 μM.

### 3.2.6 Growth curves: carbon source

An isolated colony was used to inoculate 15 ml of conventional synthetic medium. The BJ2168 strain was grown in medium that contained all of the amino acids. To keep the composition of the medium consistent, all of the amino acids were also added to the medium used to grow the *ccc2*, *fet3*, and *ctr1* strains. The cultures were grown to saturation (~30 hours) and the cells were washed with and resuspended in ice-cold, sterile, Milli-Q water. These cells were used to inoculate 5 ml of either conventional (glucose) or induction (galactose and raffinose) synthetic medium at an $OD_{600}$ of 0.1. All of the amino acids were added to both the conventional and induction

synthetic media. The amount of cells used to inoculate the various media was calculated based on measurements in a standard 1 ml cuvette (Milton Roy Spectronic 1201). All strain/media combinations were setup in triplicate. Samples (200 µl) were removed at appropriate intervals and the $OD_{600}$ was determined in a 96-well microplate (Molecular Devices SPECTRAmax).

### 3.2.7 Isolation of total RNA and Northern hybridization

Isolated *S. cerevisiae* colonies were used to inoculate 15 ml aliquots of conventional synthetic medium. The cells from stationary cultures (~30 hours) were then washed with and resuspended in ice-cold, sterile, Milli-Q water. These cells were used to inoculate 7.5 ml of induction synthetic medium at an $OD_{600}$ of 0.4. After an overnight incubation (~16 hours), the QIAgen RNeasy kit and the corresponding Yeast I Protocol (Enzymatic Lysis Protocol-Standard Version) was used to isolate total RNA. Electrophoresis, Northern transfer, and hybridization were performed as outlined in the GeneScreen™ & GeneScreen *Plus*® Hybridization Transfer Membranes, Transfer and Detection Protocols (DuPont). A 7.5 µg sample of total RNA was loaded into each lane. A *BglII* restriction fragment, isolated from the pYES3-cDNA vector (Figure 3.2), was labeled and used as the probe.

### 3.2.8 Plating assay

*Saccharomyces cerevisiae* strains that contain a disrupted *CCC2* gene are defective in high-affinity iron uptake and should be unable to grow on iron-limited medium. However, Forbes (2000) found that *ccc2* cells grew on iron-limited medium if the cells had been grown in conventional synthetic medium immediately prior to dilution and plating on iron-limited medium. This observation suggested that *S. cerevisiae* cells contain a store of intracellular iron that allows a period of normal growth when iron becomes limiting (Forbes, 2000). Growth under iron-limited conditions is, therefore, only dependent on high-affinity iron uptake after this store has been depleted. An overnight incubation in iron-limited medium was included to help deplete this store and

to allow the iron-dependent growth of the *ccc2* mutant to become apparent (Forbes, 2000).

Isolated *S. cerevisiae* colonies were used to inoculate 15 ml aliquots of conventional synthetic medium. The cultures were grown to saturation (~30 hours), washed, and resuspended in 15 ml of induction synthetic medium. After an overnight incubation (~16 hours) the cells were washed again and resuspended in 15 ml of iron-limited medium. Following another overnight incubation (~16 hours) the cells were washed once more and resuspended in ice-cold, sterile, Milli-Q water at an $OD_{600}$ of 0.1. All washes used ice-cold, sterile, Milli-Q water. The iron-limited, iron-sufficient, and copper-sufficient plates (150 mm x 15 mm) were divided into 16. A 5 µl sample of the appropriate cells was spotted onto each division and a second 5 µl sample was spread onto the same division. Each plate was setup in triplicate. Photographs were taken under sterile conditions after 72 and 240 hours. Iron-limited, iron-sufficient, and copper-sufficient plates (100 mm x 15 mm) that contained 2% glucose, and did not contain galactose and raffinose, were also divided into 16 and spotted with single 5 µl samples. These plates were incubated for 48 hours and photographed.

### 3.2.9 Growth curves: complementation assay

Isolated *S. cerevisiae* colonies were used to inoculate 15 ml aliquots of conventional synthetic medium. The cultures were grown to saturation (~30 hours), washed, and resuspended in 15 ml of induction synthetic medium. After an overnight incubation (~16 hours) the cells were washed again and resuspended in 15 ml of iron-limited medium. Following another overnight incubation (~16 hours) the cells were washed once more and resuspended in 1 ml of iron-limited medium. These cells were used to inoculate 5 ml of iron-limited, iron-sufficient, or copper-sufficient media at an $OD_{600}$ of 0.1. The amount of cells used to inoculate the various media was calculated based on measurements in a standard 1 ml cuvette (Milton Roy Spectronic 1201). All strain/media combinations were setup in triplicate. Samples (200 µl) were removed at

appropriate intervals and the $OD_{600}$ was determined in a 96-well microplate (Molecular Devices SPECTRAmax).

## 3.3 Results

### 3.3.1 Growth curves: carbon source

The BJ2168, *ccc2*, *fet3*, and *ctr1* strains showed similar growth in the conventional synthetic medium (glucose; Figure 3.3A), with fast growth rates and high final cell densities. In contrast, only the BJ2168 strain grew rapidly and to a relatively high cell density in the induction synthetic medium (galactose and raffinose; Figure 3.3B). Growth rates and final cell densities of the remaining strains was inhibited in the induction synthetic medium (galactose and raffinose; Figure 3.3B), with the greatest reduction occurring for the *fet3* strain.

The complementation assay was, therefore, initiated by first growing stationary cultures in a conventional, glucose-based, synthetic medium. The cells were then washed to help reduce the carry over of glucose (which represses transcription from the *GAL1* promoter) resuspended in synthetic induction medium, and incubated overnight to allow transcription of the *B. napus* cDNA. A similar strategy was used to obtain cultures that were actively growing and transcribing the *B. napus* cDNA, and were, consequently, suitable for RNA isolation. Since the galactose- and raffinose-based induction synthetic medium (Figure 3.3B) supported slower growth rates than the glucose-based conventional synthetic medium (Figure 3.3A), it was anticipated that the methods used by Forbes and Cox (1998), particularly the growth periods, might require further modification.

### 3.3.2 Plating assay

The iron-limited, iron-sufficient, and copper-sufficient plates, which were made with galactose and raffinose as a carbon source, were photographed after 72 hours. This incubation time was increased from 48 hours (Forbes and Cox, 1998) since the *S. cerevisiae* strains that were used in this study grow slower when galactose and raffinose replace glucose as the carbon source (Figures 3.3A and 3.3B).

The growth pattern observed on iron-limited plates (Figure 3.4B) suggested that the putative $Cu^{2+}$-ATPase encoded by the *B. napus* cDNA is capable of complementing the *ccc2* mutant. The *ccc2* strain that was transformed with the pYES3-cDNA (or pG3-ATP7B) vector was able to grow on the iron-limited medium, while the *ccc2* strain that was transformed with empty pYES3 (or pG3) vector did not grow. These results suggest that the cDNA was required for the growth of the *ccc2*/pYES3-cDNA strain on the iron-limited plate. The only *ccc2* transformant that grew on the glucose-based, iron-limited plates (data not shown) was the *ccc2*/pG3-ATP7B strain. This observation strengthens the argument that galactose-mediated transcription of the *B. napus* cDNA was required for complementation of the *ccc2* mutant by the pYES3-cDNA vector.

The *ccc2* mutant is unable to grow on the iron-limited plates as a result of a deficiency in high-affinity iron uptake. When the *ccc2* gene is disrupted, Fet3p does not receive the copper it requires to function and this impaired Fet3p activity results in the deficiency in high-affinity iron uptake. It is, therefore, possible that the *B. napus* cDNA rescued the *ccc2* mutant by functionally replacing the inactive Fet3p or by providing iron uptake activity. A *fet3* mutant strain was, consequently, transformed with all of the vectors. All of the *fet3* transformants were unable to grow on the iron-limited plates (Figure 3.4B). These results show that the *B. napus* cDNA (or *ATP7B* cDNA) was not capable of directly replacing an inactive Fet3p or a defective high-affinity iron uptake system. This provides further support for the proposal that the growth of the *ccc2* mutant carrying the pYES3-cDNA (or pG3-ATP7B) vector resulted from a direct

complementation of the *ccc2* mutation and that the product encoded by the cDNA is a functional homolog of Ccc2p.

The *ctr1* mutant is also deficient in high-affinity iron uptake. This deficiency is secondary to a defect in high-affinity copper transport and occurs because copper is not delivered to Ccc2p and Fet3p. The *ctr1* mutant was transformed with all of the vectors. All of the transformants were unable to grow on the iron-limited plates (Figure 3.4B). These results suggest that the protein encoded by the *B. napus* cDNA is unable to rescue the *ctr1* mutant and that the protein encoded by the *B. napus* cDNA is not able to mediate copper uptake across the plasma membrane.

The results of Northern hybridizations confirm that the *B. napus* cDNA was being transcribed in all of the strains that contained the pYES3-cDNA vector, but not in the strains that contained the empty pYES3 vector (Figure 3.5). This confirmation strengthens the argument that the protein encoded by the *B. napus* cDNA is responsible for growth of the *ccc2*/pYES3-cDNA strain on the iron-limited plates. It should, however, be noted that the amount of transcripts varied between the strains. The BJ2168/pYES3-cDNA strain produced a greater quantity of transcripts than the *ccc2*/pYES3-cDNA strain, while the *ctr1*/pYES3-cDNA and *fet3*/pYES-cDNA strains produced fewer transcripts than the *ccc2*/pYES3-cDNA strain (Figure 3.5). It might, therefore, be argued that the lack of complementation observed in the *ctr1*/pYES3-cDNA and *fet3*/pYES3-cDNA strains on the iron-limited plates could be due to insufficient transcript levels that correspond to the *B. napus* cDNA. This seems unlikely as the pYES3-cDNA vector is present at high-copy numbers and the *GAL1* promoter, when induced, mediates a high level of expression.

Previous studies have demonstrated that the *ccc2*, *fet3*, and *ctr1* mutants can all be rescued by supplementing an iron-limited medium with copper and/or iron (Section 1.10). When the iron-limited medium was supplemented to 500 μM copper sulfate (Figure 3.4C) all of the *ccc2* strains grew, irrespective of the vector with which they had been transformed. This demonstrates that the *ccc2*-pYES3 (or *ccc2*-pG3) strain was viable and

further supports that it was the absence of the *B. napus* cDNA (or *ATP7B* cDNA) in this strain that prevented its growth on the iron-limited medium. Copper-mediated growth rescue of the *ccc2* mutant likely resulted from Ccc2p-independent copper loading of Fet3p at the plasma membrane (Yuan *et al.*, 1997). Growth rescue was also observed when the *ctr1* mutants were grown on copper-sufficient medium (Figure 3.4C). This rescue demonstrates that these mutants were viable and, therefore, strengthens the argument that the *B. napus* cDNA (or ATP7B cDNA) was unable to rescue the *ctr1* mutants. Copper loading of Fet3p at the plasma membrane combined with low affinity uptake of copper likely played a role in copper-mediated rescue of the globally copper-deficient *ctr1* mutants. When the iron-limited medium was supplemented to 350 μM ferrous ammonium sulfate (Figure 3.4D) all of the *ccc2* strains grew. This demonstrates again that the *ccc2*-pYES3 (or *ccc2*-pG3) strain was viable. The low-affinity uptake system encoded by the *FET4* gene is credited with allowing the *ccc2* mutant to grow on iron-sufficient media. A similar iron-mediated growth rescue has been observed for *fet3* mutants (Yuan *et al.*, 1995).

In contrast to my expectations, the *fet3* strains did not grow on the iron-sufficient plates (galactose and raffinose; Figure 3.4D). The same *fet3* strains did, however, grow on glucose-based, iron-sufficient plates (data not shown). Growth of the *fet3* mutants on glucose suggested that these mutants were viable when they were plated and could be rescued by iron supplementation. A possible explanation for the lack of growth on the iron-sufficient plates that were made with galactose and raffinose, might be the slower growth rate exhibited by *fet3* mutants when galactose and raffinose are provided as the carbon source. The plates were, therefore, incubated for a longer period and with time the *fet3* mutants began to grow on the iron-sufficient media (Figure 3.6D), demonstrating that these mutants were indeed viable and that iron supplementation could rescue *fet3* mutants. It is important to note that this extended growth period did not alter the results of the iron-limited plates (Figure 3.6B). Although the *ccc2*, *fet3*, and *ctr1* mutants are all deficient in high-affinity iron uptake, the protein encoded by the *B. napus* cDNA was only capable of complementing the *ccc2* mutant. This supports a copper-transporting role

that is localized to a secretory compartment for the product encoded by the *B. napus* cDNA.

After the extended incubation, the *ctr1* mutants (which are deficient in both high-affinity copper and iron uptake) also grew on the iron-sufficient plates (Figure 3.6D). A limited amount of growth was also apparent for the *ctr1* mutants on the iron-sufficient media after short incubation (Figure 3.4D). Growth of the *ctr1* mutants was not observed on iron-limited plates after the short or extended incubation (Figures 3.4B and 3.6B). It is, therefore, likely that the major factor limiting the growth of the *ctr1* mutants on the iron-limited plates was iron-deficiency and that the high concentration of iron in the iron-sufficient plates alleviated this iron-deficiency. Under conditions such as a copper deficiency (or the presence of a non-fermentable carbon source) where the deficiency in high-affinity iron uptake is not the major determinant limiting growth of the *ctr1* mutant, iron supplementation of the media does not rescue growth of the *ctr1* mutant (Dancis *et al.*, 1994; Yuan *et al.*, 1995). None of the media used in this study were severely copper-limited and the amount of copper sulfate added to the iron-limited and iron-sufficient medium has been shown to support limited growth of *ctr1* mutants when iron was not limited (Dancis *et al.*, 1994). The *ccc2*, *fet3*, and *ctr1* mutant strains were, consequently, similar in that the major factor limiting their growth on the iron-limited medium was associated with a loss of Fet3p activity and high-affinity iron uptake. The lack of complementation of the *ctr1* mutant by transformation with the pYES3-cDNA vector, therefore, provides additional support suggesting that the *B. napus* cDNA does not encode a protein capable of directly replacing the inactive Fet3p or the defective high-affinity iron uptake system. This provides further support for the proposal that the growth of the *ccc2* mutant carrying the pYES3-cDNA (or pG3-ATP7B) vector resulted from a direct complementation of the *ccc2* mutation and that the product encoded by the *B. napus* cDNA is a functional homolog of Ccc2p.

After extended incubation, *fet3* mutants also exhibited a small amount of growth on the copper-sufficient plates (Figure 3.6C). This growth became apparent later than the growth of the *fet3* mutants on iron-sufficient plates and did not reach the same cell

density. None of the *fet3* strains grew on the iron-limited media (Figure 3.6B). These results, therefore, suggest that a mechanism may exist whereby copper is able to partially restore growth of *fet3* mutants. More importantly, they further demonstrate that the *fet3* strains were viable.

### 3.3.3 Growth curves: complementation assay

Forbes and Cox (1998) used time points of 0, 3, 6, 12, and 24 hours to create their growth curves. The media used by Forbes and Cox (1998) contained glucose, while the current study used media that replaced the glucose with galactose and raffinose. A slower growth rate was observed when galactose and raffinose were provided as the carbon source instead of glucose (Figures 3.3B and 3.3C). It was, therefore, not surprising that the strains in the iron-limited, iron-sufficient, and copper-sufficient media had not reached the stationary phase by 24 hours (Figure 3.7). In order to obtain appropriate growth curves samples were taken at 0, 12, 16, 20, 24, 36, 40, 44, and 48 hours.

The results of the growth data further support the hypothesis that the *B. napus* cDNA encodes a copper-transporting protein that is capable of functionally replacing Ccc2p. The *ccc2* mutants that were transformed with the pYES3-cDNA vector were able to grow in iron-limited medium (Figure 3.7A). Growth in the iron-limited medium was not observed for the *ccc2* strains that were transformed with the empty pYES3 vector (Figure 3.7A). These results suggest that the cDNA was responsible for the growth of the *ccc2*/pYES3-cDNA strain in the iron-limited medium. The *ccc2* strain that had been transformed with the pYES3 vector grew in both iron- and copper-sufficient media (Figures 3.7B and 3.7C), suggesting that the inoculum used to initiate all of the *ccc2*/pYES3 growth curves, including the iron-limited medium, contained viable cells.

The BJ2168/pYES3, BJ2168/pYES3-cDNA, and *ccc2*/pYES3-cDNA strains grew in a similar manner, relative to one another, in the iron-limited, iron-sufficient, and copper-sufficient media (Figure 3.7). These results indicate that the *B. napus* cDNA-

mediated rescue (*ccc2*/pYES3-cDNA) resulted in copper delivery to Fet3p that was at least equivalent to the amount of copper supplied to Fet3p by the endogenous *CCC2* gene (BJ2168/pYES3). It should be noted that this does not imply that an individual protein encoded by the *B. napus* cDNA would exhibit the same level of activity as a single Ccc2 protein. The number of proteins contributing to copper delivery to Fet3p likely varies in the *ccc2*/pYES3-cDNA strain and the BJ2168 strains. This difference in protein number is expected since the transcription of the genes is under the control of different promoters and the *CCC2* gene is present as a single copy, while the cDNA is present as multiple copies. As evident from the growth curves of the BJ2168/pYES3 and BJ2168/pYES3-cDNA strains (Figures 3.7A, 3.7B, and 3.7C), the protein encoded by the *B. napus* cDNA was neither detrimental nor beneficial, under the conditions tested, when the functional intracellular copper-transporter Ccc2p is already present.

## 3.4 Discussion

A *B. napus* cDNA encoding a putative $P_{1B}$-ATPase was cloned in this study. An analysis of the amino acid sequence that was predicted from this cDNA suggested that the putative $P_{1B}$-ATPase was likely a copper transport protein. A complementation assay was used to demonstrate that the protein encoded by the *B. napus* cDNA is able to functionally replace the $Cu^{2+}$-ATPase Ccc2p from *S. cerevisiae*. The *B. napus* protein was able to rescue growth of a *ccc2* mutant under iron-limited conditions. Although copper transport remains to be demonstrated directly, this rescue presumably resulted from a restoration of copper delivery to Fet3p that was mediated by the protein encoded by the *B. napus* cDNA.

During the progress of this research, an *Arabidopsis* cDNA corresponding to the genomic locus T19K24.18 (*RAN1*) from the *RAN1*/AC002342/2660661 genomic record, was cloned (Hirayama *et al.*, 1999). This cDNA was cloned as a result of a search for the mutant gene responsible for an alteration in the specificity of the ethylene receptor(s). Exposure of dark-grown *Arabidopsis* seedlings to exogenous ethylene results in the ethylene triple response; inhibition of root and hypocotyl elongation, radial swelling of

the hypocotyl, and exaggerated apical hook growth. *Trans*-cyclooctene (TCO) is an ethylene receptor antagonist (a potent competitive inhibitor of ethylene binding to its receptor(s)) that prevents the ethylene-induced triple response in wild-type *Arabidopsis* seedlings. Hirayama *et al.* (1999) identified two allelic *Arabidopsis* mutants that displayed the "ethylene" triple response when treated with TCO. They, consequently, named these mutants *responsive-to-antagonist1* (*ran1*). Hirayama *et al.* (1999) used positional cloning to narrow the *RAN1* locus to the BAC clone T19K24 (AC002342/2660661). Since ethylene binding to its receptor has been proposed to require a transition metal, this gene was considered as a candidate for *RAN1*. After single base changes were observed in the two mutant alleles, *ran1-1* and *ran1-2*, the cDNA corresponding to the T19K24.18 genomic locus was cloned.

The RAN1 protein has been proposed to function in the delivery of copper to ethylene receptor(s) (Hirayama *et al.*, 1999; Woeste and Kieber, 2000). The ethylene receptor(s) are negative regulators that, when active, function to prevent the ethylene response phenotypes in the absence of the hormone. Binding of ethylene (or TCO in the *ran1-1* and *ran1-2* mutants) by the ethylene receptor(s) is thought to inactivate the receptor(s). Inactivation of the receptor(s) results in derepression of the downstream signaling pathway components and activation of the hormone response phenotypes. The *ran1-1* and *ran1-2* agonist-like response to TCO has been proposed to be a result of suboptimal copper:apoprotein stoichiometry that might alter the conformation of the receptor(s) and allow reduced ligand specificity. *Arabidopsis* plants that display a reduced level of *RAN1* mRNA (Hirayama *et al.*, 1999) or a strong loss of function mutation (*ran1-3*; Woeste and Kieber, 2000) display a constitutive ethylene response that is thought to be due to copper-deficient, and consequently inactive, ethylene receptor(s). Reminiscent of the copper-mediated rescue of *ccc2* mutants, addition of copper ions to the plant growth medium partially suppresses the phenotypes of the *ran1-1*, *ran1-2*, and *ran1-3* mutants.

The RAN1 protein may also deliver copper to proteins other than the ethylene receptors (Woeste and Kieber, 2000). The RAN1 protein functions early in the ethylene-

signaling pathway (Hirayama *et al.*, 1999) and mutants that are ethylene-insensitive have been identified downstream of RAN1. Aspects of the *ran1-3* mutant phenotype that are not suppressed by the ethylene-insensitive mutations have been proposed to result from an ethylene-independent pathway. While the ethylene-insensitive mutations suppressed the constitutive ethylene phenotypes, the rosette lethal phenotype was not suppressed. The rosette lethal phenotype has been proposed to arise from an ethylene-independent pathway regulating cell expansion that is dependent upon cuproenzyme(s) that receive copper in a RAN1-dependent manner and are, consequently, inactive in the *ran1-3* mutant.

The *B. napus* and *RAN1* coding sequences are 88% identical, while the predicted amino acid sequences are 91% identical. This level of sequence identity suggests that the *B. napus* cDNA encodes a RAN1 homologue. It is, consequently, likely that the protein encoded by the *B. napus* cDNA mediates a function similar to that described for RAN1. Indeed, the results of the complementation assay suggest that the *B. napus* cDNA encodes a copper-transporting function and the mutations identified in the *ran1-1*, *ran1-2*, and *ran1-3* mutants are not found in the *B. napus* cDNA.

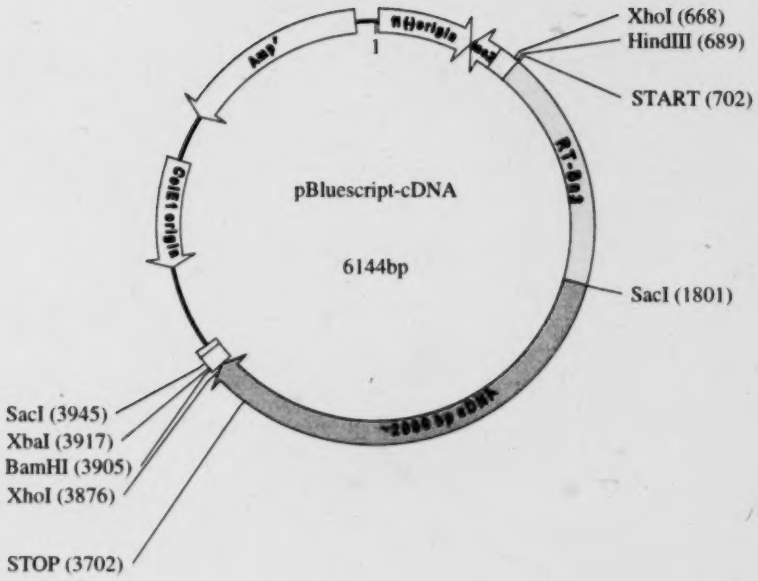


**Figure 3.1 The full-length *B. napus* cDNA in the pBluescript SK- plasmid (pBluescript-cDNA).** The pBluescript-cDNA was created by cloning the RT-Bn3 (or RT-Bn2) RT-PCR fragment with *XhoI / SacI* ends (light grey) and the ~2000 bp cDNA fragment with *SacI / XbaI* ends (dark grey) into a *XhoI / XbaI* digested pBluescript SK- vector (white).

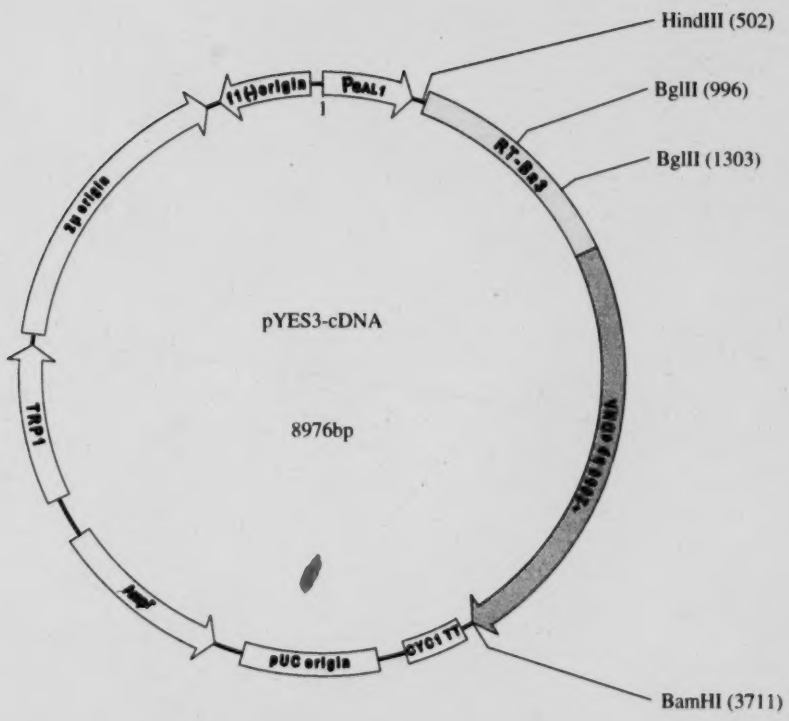


**Figure 3.2 The full-length *B. napus* cDNA in the pYES3 vector (pYES3-cDNA).** The full-length cDNA that was liberated from the pBluescript SK- vector (Figure 3.1) using the *HindIII* and *BamHI* restriction enzymes was cloned into a *HindIII* / *BamHI*-digested pYES3 vector. The *BglII* fragment (307 bp; position 996 to 1303) that contained only cDNA sequence was used to produce the probe for the Northern hybridizations.

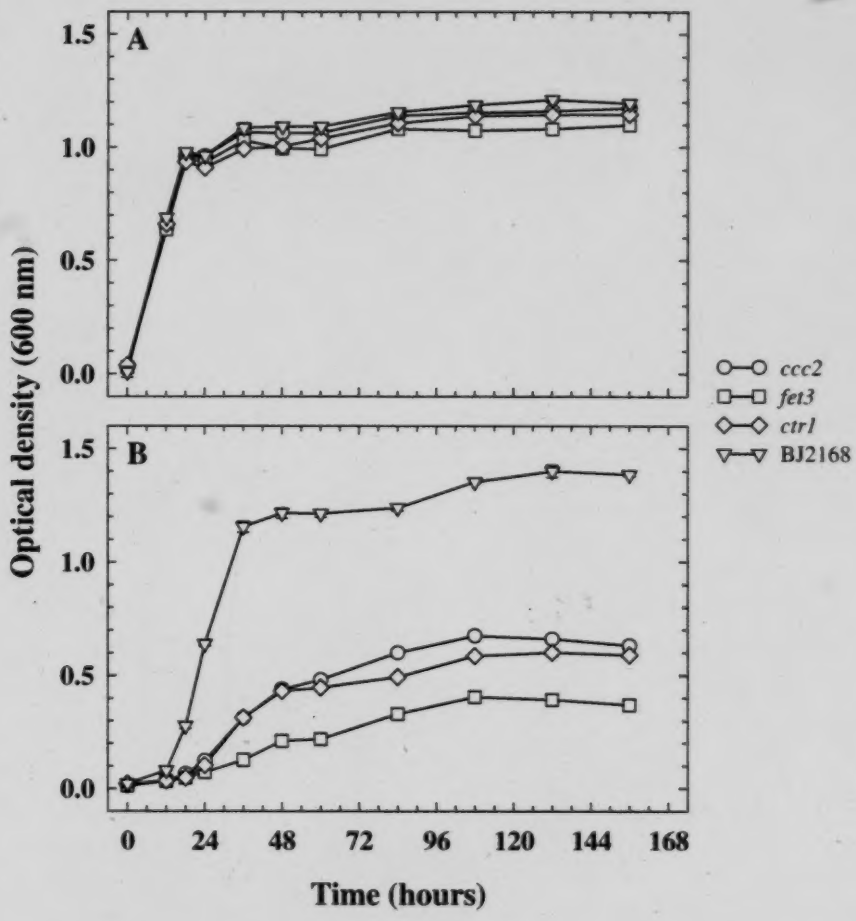


**Figure 3.3 Carbon source growth curves.** Cultures were grown to saturation in conventional synthetic medium. Cells were washed, resuspended in ice-cold, sterile, Milli-Q water, and used to inoculate 5 ml of conventional or induction synthetic media at an $OD_{600}$ of 0.1 (as determined using a standard 1 ml cuvette and the Milton Roy Spectronic 1201). Each data point represents the average $OD_{600}$ (n=3; ±SE), as measured in a 96 well microplate reader (Molecular Devices SPECTRAmax spectrophotometer). The bars represent the standard errors. Repetition of the experiment produced results similar to those shown. **A.** Conventional synthetic media (glucose as a carbon source). **B.** Induction synthetic media (galactose and raffinose as a carbon source).

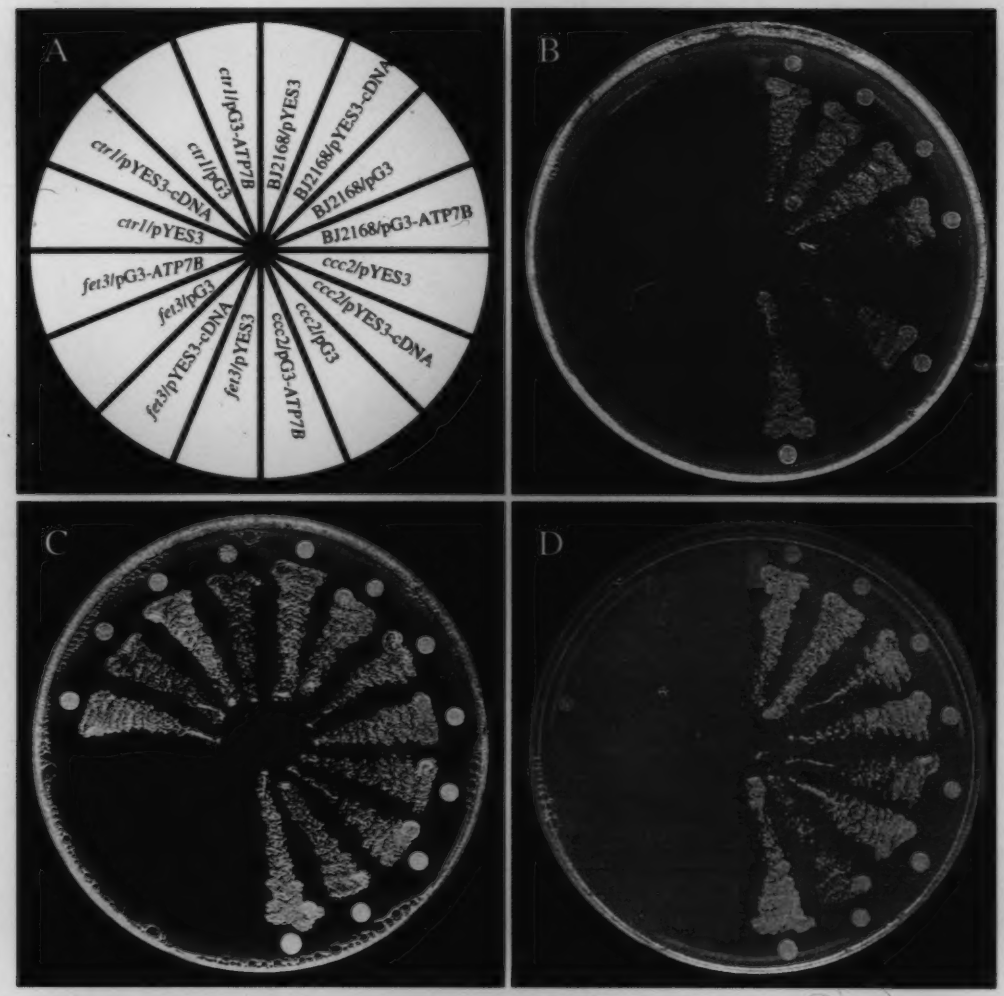


**Figure 3.4 Complementation assay plates after a short (72 hour) incubation.** Stationary cells were grown, from isolated colonies, in conventional synthetic medium. These cells were washed, resuspended in induction synthetic medium, incubated, washed again, resuspended in iron-limited medium, and incubated again. Cells were then resuspended at an $OD_{600}$ of 0.1 in ice-cold, sterile, Milli-Q water and 5 µl samples of these resuspended cells were spotted or spread onto the plates that were then incubated at 30°C for 72 hours. The plates shown represent one of three plates that were set-up in a single experiment; the results of the two remaining plates were consistent with those that are shown. Repetition of the experiment produced results comparable to those shown. **A.** Pattern used to setup the plates. **B.** Iron-limited plate. **C.** Copper-sufficient plate. **D.** Iron-sufficient plate.

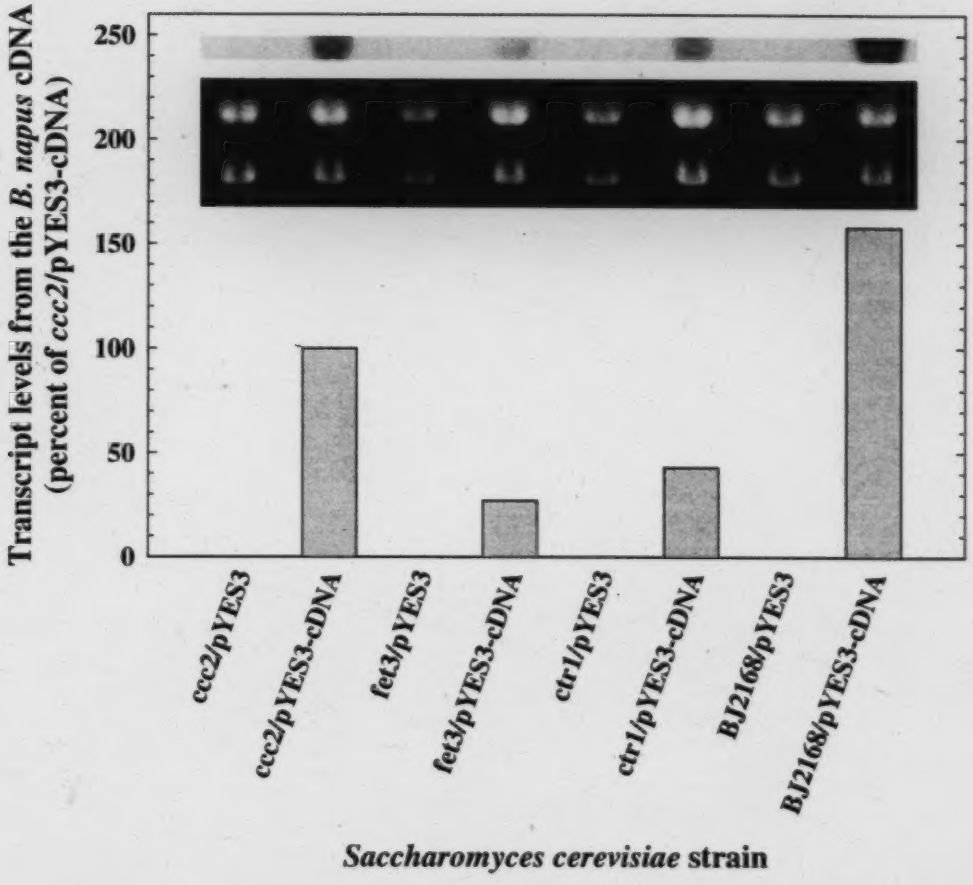


**Figure 3.5 Transcript levels from the *B. napus* cDNA.** Total RNA was isolated from *S. cerevisiae* strains using the QIAgen RNeasy kit (Section 3.2.7). Cultures that were used for total RNA isolation were first grown in conventional synthetic medium and were later incubated in induction synthetic medium (Section 3.2.7). Electrophoresis, Northern transfer, and hybridization were performed as outlined in the GeneScreen™ & GeneScreen *Plus*® Hybridization Transfer Membranes, Transfer and Detection Protocols (DuPont). A 7.5 µg sample of total RNA was loaded into each lane. A *BglII* restriction fragment, isolated from the pYES3-cDNA vector (Figure 3.2), was labeled and used as the probe. The results of the Northern hybridization and the ribosomal RNA from the gel that was used in the Northern transfer are inset in the graph.

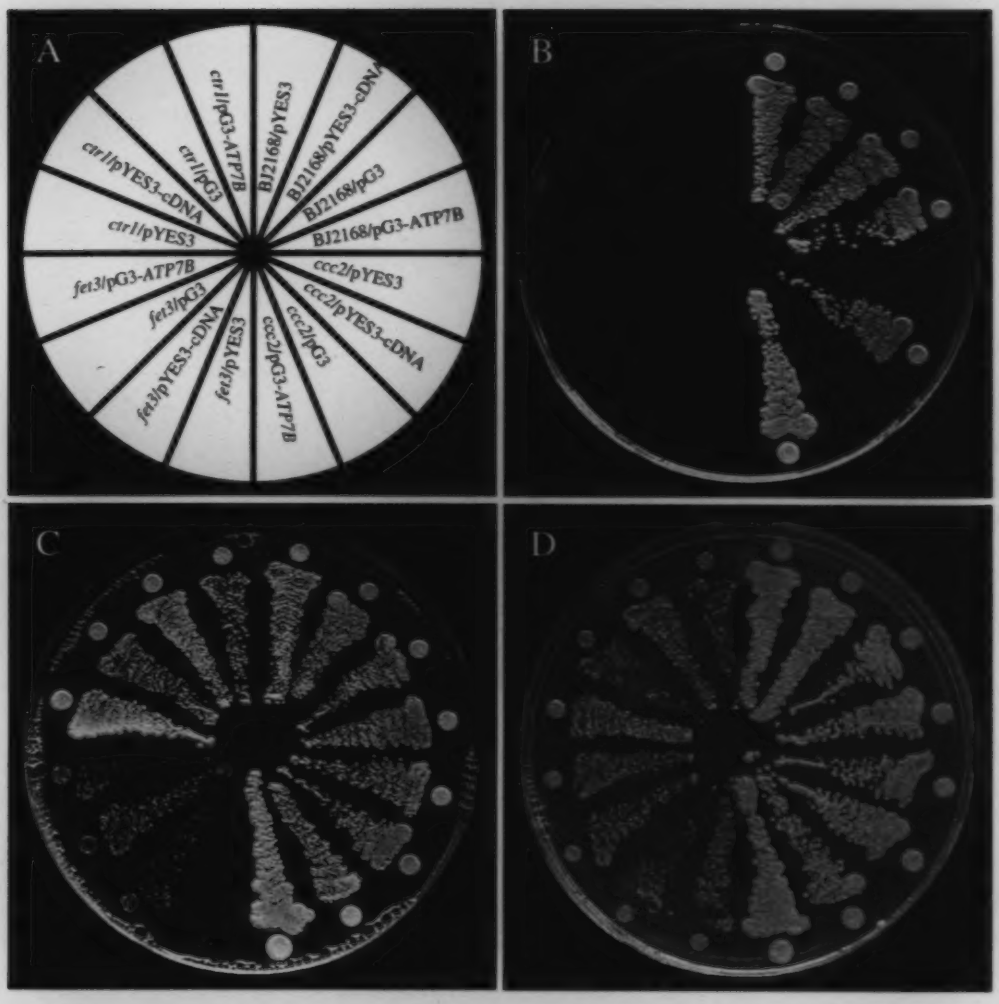


**Figure 3.6 Complementation assay plates after an extended (240 hour) incubation.** Stationary cells were grown, from isolated colonies, in conventional synthetic medium. These cells were washed, resuspended in induction synthetic medium, incubated, washed again, resuspended in iron-limited medium, and incubated again. Cells were then resuspended at an $OD_{600}$ of 0.1 in ice-cold, sterile, Milli-Q water and 5 µl samples of these resuspended cells were spotted or spread onto the plates that were then incubated at 30°C for 240 hours. The plates shown represent one of three plates that were set-up in a single experiment; the results of the two remaining plates were consistent with those that are shown. Repetition of the experiment produced results comparable to those shown. **A.** Pattern used to setup the plates. **B.** Iron-limited plate. **C.** Copper-sufficient plate. **D.** Iron-sufficient plate.

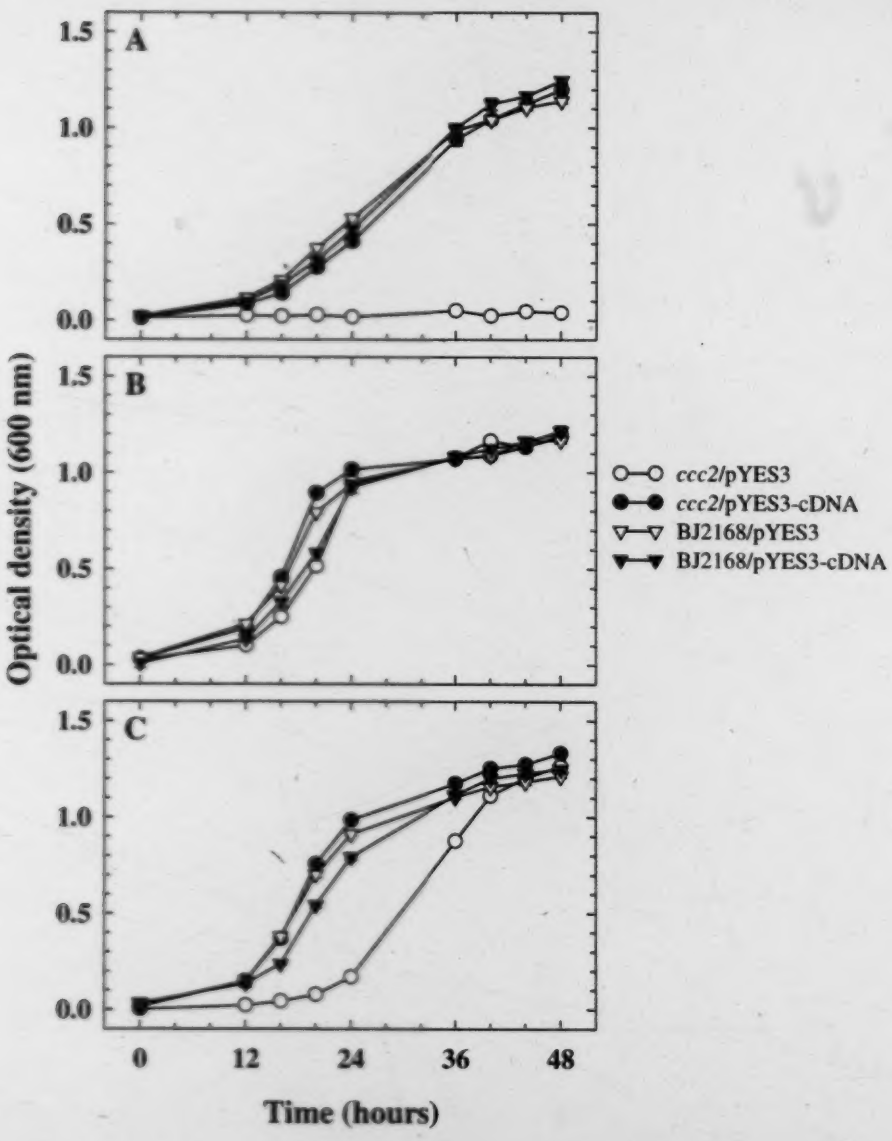


**Figure 3.7 Complementation assay growth curves.** Stationary cultures were grown in conventional synthetic medium. Cells from these cultures were washed, resuspended in induction synthetic medium, incubated, washed again, resuspended in iron-limited medium, and incubated again. Cells were then resuspended in iron-limited medium and used to inoculate 5 ml of the appropriate media at an $OD_{600}$ of 0.1 (as determined using a standard 1 ml cuvette and the Milton Roy Spectronic 1201). Each data point represents the average $OD_{600}$ (n=3; ±SE), as measured in a 96 well microplate reader (Molecular Devices SPECTRAmax spectrophotometer). Repetition of the experiment produced results similar to those shown. **A.** Iron-limited medium. **B.** Copper-sufficient medium. **C.** Iron-sufficient medium.

## 3.5 Literature cited

**Dancis A, Haile D, Yuan DS, Klausner RD** (1994) The *Saccharomyces cerevisiae* copper transport protein (Ctr1p). Biochemical characterization, regulation by copper, and physiologic role in copper uptake. J Biol Chem **269(41)**:25660-25667

**Davis-Kaplan SR, Askwith CC, Bengtzen AC, Radisky D, Kaplan J** (1998) Chloride is an allosteric effector of copper assembly for the yeast multicopper oxidase Fet3p: an unexpected role for intracellular chloride channels. Proc Natl Acad Sci U S A. **95(23)**:13641-13645

**Elble R** (1992) A simple and efficient procedure for transformation of yeasts. Biotechniques **13(1)**:18-20.

**Forbes JR** (2000) Functional analysis of the copper-transporting P-type ATPase, ATP7B, defective in Wilson disease. PhD thesis. University of Alberta, Canada, 2000 159 pp

**Forbes JR, Cox DW** (1998) Functional characterization of missense mutations in *ATP7B*: Wilson disease mutation or normal variant? Am J Hum Genet **63(6)**:1663-1674

**Hirayama T, Kieber JJ, Hirayama N, Kogan M, Guzman P, Nourizadeh S, Alonso JM, Dailey WP, Dancis A, Ecker JR** (1999) RESPONSIVE-TO-ANTAGONIST1, a Menkes/Wilson disease-related copper transporter, is required for ethylene signaling in *Arabidopsis*. Cell **97(3)**:383-393

**Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD, Gitlin JD** (1997) Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. J Biol Chem **272(34)**:21461-21466

**Iida M, Terada K, Sambongi Y, Wakabayashi T, Miura N, Koyama K, Futai M, Sugiyama T** (1998) Analysis of functional domains of Wilson disease protein (ATP7B) in *Saccharomyces cerevisiae*. FEBS Lett **428(3)**:281-285

**Payne AS, Gitlin JD** (1998) Functional expression of the Menkes disease protein reveals common biochemical mechanisms among the copper-transporting P-type ATPases. J Biol Chem **273(6)**:3765-3770

**Sambongi Y, Wakabayashi T, Yoshimizu T, Omote H, Oka T, Futai M** (1997) *Caenorhabditis elegans* cDNA for a Menkes/Wilson disease gene homologue and its function in a yeast *CCC2* gene deletion mutant. J Biochem (Tokyo) **121(6)**:1169-1175

**Schena M, Picard D, Yamamoto KR** (1991) Vectors for constitutive and inducible gene expression in yeast. Methods Enzymol **194**:389-398.

**Sherman F** (1991) Getting started with yeast. Methods Enzymol **194**:3-21

**Sherman F** (1997) Yeast genetics. In The Encyclopedia of Molecular Biology and Molecular Medicine Vol 6. (R. Meyers, ed.) VCH Publishing, Weinheim, Germany, 1997 pp. 302-325

**Woeste KE, Kieber JJ** (2000) A strong loss-of-function mutation in *RAN1* results in constitutive activation of the ethylene response pathway as well as a rosette-lethal phenotype. Plant Cell **12(3)**:443-455

**Yuan DS, Stearman R, Dancis A, Dunn T, Beeler T, Klausner RD** (1995) The Menkes/Wilson disease gene homologue in yeast provides copper to a ceruloplasmin-like oxidase required for iron uptake. Proc Natl Acad Sci U S A **92(7)**:2632-2636

**Zubenko GS, Mitchell AP, Jones EW** (1980) Mapping of the proteinase b structural gene *PRB1*, in *Saccharomyces cerevisiae* and identification of nonsense alleles within the locus. Genetics **96(1)**:137-146

# 4 Concluding discussion

A full-length *Brassica napus* cDNA that encodes a putative $P_{1B}$-ATPase was cloned in this study. The sequence of this cDNA was submitted to GenBank (AY045772/15636780) and recent BLAST searches (January, 2002) reveal that this sequence is the first and only full-length sequence available in the NCBI sequence databases that encodes a putative $P_{1B}$-ATPase from a plant other than *Arabidopsis*. The amino acid sequence (AAL02122/15636781) that was predicted from the *B. napus* cDNA was most similar to the $Cu^{2+}$-ATPases and copper was, consequently, considered to be a likely substrate. A *Saccharomyces cerevisiae* complementation assay provided evidence suggesting that the *B. napus* cDNA does encode a functional copper-transporting protein. When this study was initiated only four full-length plant sequences, *PAA1*, *RAN1*, *HMA1*, and *HMA4*, that encoded putative $P_{1B}$-ATPases were reported in the literature and/or sequence databases. Although a cDNA encoding PAA1 had been cloned (Tabata *et al.*, 1997), there was no experimental evidence available to suggest that any of these $P_{1B}$-ATPases were functional and no information regarding their substrate specificity. During the progress of this study, a paper was published that suggested that RAN1 is able to transport copper and proposed that this transport plays an important role in the formation of functional ethylene receptors (Hirayama *et al.*, 1999). The experimental information that is available for all of the other putative $P_{1B}$-ATPases from plants is still limited (see below). The current study is, therefore, amongst the first studies to demonstrate that a putative plant $P_{1B}$-ATPase is functional and to examine its substrate specificity. It should, however, be noted that the protein encoded by the *B. napus* cDNA is likely a RAN1 homologue. The results of this study further support the suggestion that RAN1 (Hirayama *et al.*, 1999) and its homologues transport copper.

The *S. cerevisiae* complementation assays that were used to study the proteins encoded by the *RAN1* (Hirayama *et al.*, 1999) and the *B. napus* (this study) cDNAs both utilized the *S. cererivisiae ccc2* mutant. Hiryama *et al.* (1999) examined Fet3p oxidase activity and high-affinity iron uptake to suggest that RAN1 was able rescue the *ccc2* mutant. Growth rescue under iron-limited conditions was used in this study to

demonstrate that the protein encoded by the *B. napus* cDNA was able to complement the *ccc2* mutant. It is possible that the restoration of oxidase activity, high-affinity iron uptake, and growth under iron-limited conditions resulted from direct complementation of the inactive apoFet3p. The inability of the protein encoded by the *B. napus* cDNA to rescue *fet3* and *ctr1* mutants suggests that this is not the case and further supports the suggestion that the *B. napus* (or *RAN1*) cDNA encodes a protein that is a functional homologue of Ccc2p.

The RAN1 protein has been proposed to function in the delivery of copper to the ethylene receptor(s) (Hirayama *et al.*, 1999; Woeste and Kieber, 2000) and other unidentified copper-requiring proteins (Woeste and Kieber, 2000). It is likely that the protein encoded by the *B. napus* cDNA functions in a manner that is similar to RAN1. The cellular location(s) of the proteins that are encoded by *RAN1* and the *B. napus* cDNA remain to be demonstrated; it is, however, likely that these proteins deliver copper to the secretory pathway. The *S. cerevisiae* $P_{1B}$-ATPase, Ccc2p, has been shown to function in a late or post-Golgi vesicle (Yuan *et al.*, 1997). The human $P_{1B}$-ATPases ATP7A and ATP7B, which are able to rescue the *ccc2* mutant and share a relatively high level of sequence identity with the proteins encoded by the *RAN1* and *B. napus* cDNAs, are localized to the *trans*-Golgi network when copper levels are low (Petris *et al.*, 1996; Yamaguchi *et al.*, 1996; Dierick *et al.*, 1997; Hung *et al.*, 1997; Schaefer *et al.*, 1999). When copper levels are increased, the ATP7A protein redistributes to the plasma membrane, where it is believed to function in copper efflux that helps to prevent excess copper accumulation (Petris *et al.*, 1996). When the copper levels are increased, the ATP7B protein redistributes to cytoplasmic vesicles, where it is believed to play an excretory role that involves membrane vesicles (Schaefer *et al.*, 1999). Once copper levels are reduced, ATP7A and ATP7B return to the *trans*-Golgi (Petris *et al.*, 1996; Schaefer *et al.*, 1999). It remains to be determined if other $P_{1B}$-ATPases, including those encoded by the *RAN1* and *B. napus* cDNAs, undergo a similar redistribution.

Although RAN1 has been studied to the greatest extent, a limited amount of information is also available for the other putative plant $P_{1B}$-ATPases. When this project

was initiated, a single plant $P_{1B}$-ATPase, named *PAA1* for *P-type ATPase of Arabidopsis*, had been cloned (Tabata *et al.*, 1997). Tabata *et al.* (1997) did not present any functional information for PAA1 and a literature search (January, 2002) revealed that there has been no additional information published for PAA1. The cDNAs for the *HMA1* (AJ400906/7710953) and *HMA4* (AJ297264/11990206) genes have been cloned. The authors that submitted the *HMA1*mRNA/CDS (AJ400906/7710953) state that transposon-insertion mutants have been isolated for the *HMA1* gene and that they intend to study these mutants to help clarify the physiological role of the HMA1 protein (http://sobs.soton.ac.uk/staff/lew.html). Other publications briefly mention that another $P_{1B}$-ATPase, named AMA1 for *Arabidopsis* heavy metal ATPase 1, has been cloned from *Arabidopsis* (Fox and Guerinot, 1998; Palmgren and Harper, 1999; Williams *et al.*, 2000). It is unclear which, if any, of the *Arabidopsis* $P_{1B}$-ATPases that are found in the databases are equivalent to AMA1. A plant line homozygous for a T-DNA knock-out of AMA1 (*ama1-1*) has been created and an analysis of heavy metal content in the resulting *ama1-1* seedlings revealed that molybdenum was the only essential micronutrient showing a reduced accumulation (Palmgren and Harper, 1999). The AMA1 protein, which was detected using two different polyclonal antibodies to N- and C-terminal domains, was enriched in the plasma membrane fraction and was found at the highest levels in the root and flowers (Williams *et al.*, 2000).

Four additional *Arabidopsis* sequences, *HMA2*, *HMA3*, *HMA5*, and "*HMA6*", that encode putative $P_{1B}$-ATPases were added to the NCBI databases during the progress of this study. Functional information is not available for these putative $P_{1B}$-ATPases and there is no evidence to suggest that cDNAs that encode HMA2, HMA3, HMA5, or "HMA6" have been cloned. The *Arabidopsis* genome, therefore, encodes at least eight putative $P_{1B}$-ATPases. This is higher than the number of putative $P_{1B}$-ATPases that have been observed in the completely sequenced genomes of other eukaryotic organisms, including *Saccharomyces cerevisiae* (two), *Drosophila melanogaster* (one), and *Caenorhabditis elegans* (one; Axelsen and Palmgren, 2001). A number of reviews that focus, in part, on some of the *Arabidopsis* $P_{1B}$-ATPases have now been published (Palmgren and Harper, 1999; Williams *et al.*, 2000; Axelsen and Palmgren, 2001), the

putative $P_{1B}$-ATPase "HMA6", which was recently added to the databases, is not, however, discussed by any of these reviews. The P-type ATPase database web site (http://www.Patbase.kvl.dk), which is maintained by Kristian B. Axelsen, provides an useful link to P-type ATPase sequences and information.

Members of the P-type ATPase superfamily cluster according to the substrates they transport (Axelsen and Palmgren, 1998) and an analysis of the amino acid sequence of a P-type ATPase may, therefore, be of use in predicting its substrate specificity and in providing preliminary information about the possible function of the P-type ATPase. The $P_{1B}$-ATPases can be further divided into at least two groups; the first group has been described as containing all known $Cu^{2+}$-ATPases, while the second group contains the $Cd^{2+}$-ATPases (Axelsen and Palmgren, 1998). The CopB protein (a $Cu^{2+}$-ATPase from *Enterococcus hirae*) was shown to transport both copper and silver in an ATP-dependent manner, silver transport was, however, considered to be fortuitous and copper was suggested to be the natural substrate (Solioz and Odermatt, 1995). It is important to note that silver transport by $P_{1B}$-ATPases may, under certain circumstances, be of interest. Silver is a non-essential, toxic metal that is used as an antimicrobial agent (Gupta and Silver, 1998). Silver-resistant bacteria have evolved and a $P_{1B}$-ATPase has been implicated, in part, in this resistance (Gupta *et al.*, 1999). The CadA protein (a $Cd^{2+}$-ATPase from *Staphylococcus aureus*) was initially shown to catalyze ATP-dependent cadmium transport (Tsai *et al.*, 1992). Further studies focusing on CadA have suggested that zinc and lead may also be transported by this $P_{1B}$-ATPase (Rensing *et al.*, 1998) and another study has suggested that cobalt may also be transported by $P_{1B}$-ATPases that cluster with the $Cd^{2+}$-ATPases (Rutherford *et al.*, 1999). The possible substrate specificity of the proteins that cluster with CadA might, consequently, be described better as cadmium, zinc, lead, and/or cobalt. The evidence for the transport of cadmium, zinc, lead, and cobalt by $P_{1B}$-ATPases is from studies that have focused on prokaryotic proteins. It remains to be determined if eukaryotic $P_{1B}$-ATPases are also able to transport cadmium, zinc, lead, and/or cobalt and if the transport of these substrates is physiologically relevant in eukaryotic cells.

The *Arabidopsis* HMA1, HMA2, HMA3, and HMA4 proteins are most similar to the $Cd^{2+}/Zn^{2+}/Pb^{2+}/Co^{2+}$-ATPases and might, therefore, be involved in the transport of these ions. Zinc is an essential micronutrient that is toxic when present in excess. Mechanisms are, therefore, required to obtain zinc, to deliver it to appropriate locations or proteins, and to prevent its overaccumulation and toxicity. Lead, cadmium, and cobalt are non-essential, toxic metals and mechanisms are required to prevent toxicity. The *Arabidopsis* putative $P_{1B}$-ATPases HMA1, HMA2, HMA3, and HMA4 are the first eukaryotic $P_{1B}$-ATPases that share a higher level of sequence identity with the $Cd^{2+}/Zn^{2+}/Pb^{2+}/Co^{2+}$-ATPases than the $Cu^{2+}/Ag^{2+}$-ATPases (Axelsen and Palmgren, 2001). The HMA2, HMA3, and HMA4 amino acid sequences differ from other previously reported $P_{1B}$-ATPases since they appear to contain a metal-binding region in the C-terminus instead of the N-terminus. These sequences may, therefore, represent another group within the Type IB subfamily that has yet to be identified and might exhibit a substrate specificity that has not been observed for other $P_{1B}$-ATPases. The HMA2, HMA3, and HMA4 sequences were not available when the last large-scale phylogenetic analysis of the P-type superfamily was performed (Axelsen and Palmgren, 1998). The HMA1 protein is also unique in that it contains the SPC tripeptide in place of the CPx motif. This SPC motif was also observed in other putative $P_{1B}$-ATPases and together these sequences might represent yet another group within the Type IB subfamily. It will be interesting to learn the substrate specificity, subcellular location, and ultimately the function of HMA1, HMA2, HMA3, and HMA4.

The putative $P_{1B}$-ATPases encoded by *PAA1*, *RAN1*, *HMA5*, "*HMA6*", and the *B. napus* cDNA are most similar to the $Cu^{2+}/Ag^{2+}$-ATPases and are, therefore, likely to be involved in copper transport. Evidence has been presented to suggest that *RAN1* (Hirayama *et al.*, 1999) and the *B. napus* cDNA (this study) do, in fact, encode proteins that are able to transport copper. Additional observations strengthen the suggestion that PAA1, HMA5, and "HMA6" might also transport copper. The amino acid residue that is located 21 amino acids C-terminal of the CxxC motif(s) may modulate metal-binding affinity/specificity (Gitschier *et al.*, 1998). A leucine (L) residue is present 21 amino acids C-terminal of the CxxC motif(s) that are found in many $Cu^{2+}$-ATPases, including

CopA, Ccc2p, ATP7A, and ATP7B (Gitschier *et al.*, 1998). A leucine residue is also found 21 amino acids C-terminal of the CxxC motif(s) that are encoded by *PAA1*, *RAN1*, and the *B. napus* cDNA cloned in this study. The HMA5 protein, which contains two CxxC motifs, contains a leucine residue 22 amino acids C-terminal of the first CxxC motif and another leucine 21 amino acids C-terminal of the second CxxC motif. The "HMA6" protein contains a leucine residue 22 amino acids C-terminal of its CxxC motif. An extended CPx motif (CPC[AS]LGLATP) has also been proposed for $P_{1B}$-ATPases that transport copper (Tottey *et al.*, 2001). The proteins that are encoded by *PAA1*, *RAN1*, *HMA5*, "*HMA6*", and the *B. napus* cDNA cloned in this study all contain the extended CPx motif CPCALGLATP. Although an analysis of the amino acid sequences of PAA1, HMA5, and "HMA6" suggest that these putative $P_{1B}$-ATPases are able to transport copper, it is important to verify this observation experimentally.

The function of PAA1, HMA5, and "HMA6" remains to be determined. An important step in determining the function of PAA1, HMA5, and "HMA6" may be to identify where these proteins are localized. Studies in other organisms suggest that $P_{1B}$-ATPases may be localized to the plasma membrane (Petris *et al.*, 1996) and/or intracellular membranes (Kanamaru *et al*,. 1994; Hung *et al.*, 1997; Yuan *et al.*, 1997; Schaefer *et al.*, 1999; Petris *et al.*, 1996; Dierick *et al.*, 1997; Yamaguchi *et al.*, 1996). The $P_{1B}$-ATPases that are localized to the plasma membrane may be involved in the uptake or efflux of copper (Odermatt *et al.*, 1993; Petris *et al.*, 1996). The $P_{1B}$-ATPases that are localized to intracellular membranes may be responsible for compartmentalization of copper, for incorporation into proteins (Hung *et al.*, 1997; Yuan *et al.*, 1997; Schaefer *et al.*, 1999; Petris *et al.*, 1996; Dierick *et al.*, 1997; Yamaguchi *et al.*, 1996), and/or for prevention of toxicity (Schaefer *et al.*, 1999). Numerous studies have proposed a role for eukaryotic $P_{1B}$-ATPases in the delivery of copper to the secretory system (Hung *et al.*, 1997; Yuan *et al.*, 1997; Schaefer *et al.*, 1999; Petris *et al.*, 1996; Dierick *et al.*, 1997; Yamaguchi *et al.*, 1996). Studies that have focused on the $P_{1B}$-ATPase, ATP7B, suggest that in addition to delivering copper to the secretory system, ATP7B may play an excretory role when it is localized to cytoplasmic vesicles (Schaefer *et al.*, 1999). Other studies suggest that $P_{1B}$-ATPases may play a role in copper

delivery within the chloroplast (Kanamaru *et al*,. 1994; Tottey *et al*., 2001). Cyanobacteria are prokaryotic microorganisms that contain a photosynthetic apparatus that resembles that of plants. Some species of cyanobacteria have, consequently, been used as model organisms to investigate the structure and function of plant-type photosynthesis. A $P_{1B}$-ATPase, PacS, in *Synechococcus* strain 7942 is localized to the thylakoid membrane (Kanamaru *et al*,. 1994) and studies suggest that a PacS homologue in *Syechocystis* PCC6803 may be required for the delivery of copper to plastocyanin, an electron carrier that shuttles electrons between photosystems II and I (Tottey *et al*., 2001).

The $P_{1B}$-ATPases, which were once thought to be limited to bacterial species (Solioz and Vulpe, 1996), are emerging as an important group of proteins in a variety of organisms including the model plant *Arabidopsis*. The cloning of a *B. napus* cDNA that encodes a putative $P_{1B}$-ATPase that is able to transport copper (this study) supports the presence of $P_{1B}$-ATPases in plants other than *Arabidopsis*. Recent BLAST searches (January, 2002) reveal that this sequence is the first and only full-length sequence available in the NCBI sequence databases that encodes a putative $P_{1B}$-ATPase from a plant other than *Arabidopsis*. The identification of partial sequences and expressed sequence tags (ESTs) that appear to correspond to other putative $P_{1B}$-ATPases suggests that these ATPases occur in a variety of plant species. As demonstrated in this study, analysis of the amino acid sequences of $P_{1B}$-ATPases might aid in the prediction of substrate specificity. It will be interesting to learn the substrate specificity of the other putative *Arabidopsis* $P_{1B}$-ATPases that have been identified. It will also be interesting to learn the function of $P_{1B}$-ATPases in a variety of plant species.

## 4.1 Literature cited

**Axelsen KB, Palmgren MG** (1998) Evolution of substrate specificities in the P-type ATPase superfamily. J Mol Evol **46(1)**:84-101

**Axelsen KB, Palmgren MG** (2001) Inventory of the superfamily of P-type ion pumps in *Arabidopsis*. Plant Physiol **126(2)**:696-706

**Dierick HA, Adam AN, Escara-Wilke JF, Glover TW** (1997) Immunocytochemical localization of the Menkes copper transport protein (ATP7A) to the trans-Golgi network. Hum Mol Genet **6(3)**:409-416

**Fox TC, Guerinot ML** (1998) Molecular biology of cation transport in plants. Annu Rev Plant Physiol Plant Mol Biol **49**:669-696

**Gitschier J, Moffat B, Reilly D, Wood WI, Fairbrother WJ** (1998) Solution structure of the fourth metal-binding domain from the Menkes copper-transporting ATPase. Nat Struct Biol **5(1)**:47-54

**Gupta A, Matsui K, Lo JF, Silver S** (1999) Molecular basis for resistance to silver cations in *Salmonella*. Nat Med **5(2)**:183-188

**Gupta A, Silver S** (1998) Silver as a biocide: will resistance become a problem? Nat Biotechnol **16(10)**:888

**Hirayama T, Kieber JJ, Hirayama N, Kogan M, Guzman P, Nourizadeh S, Alonso JM, Dailey WP, Dancis A, Ecker JR** (1999) RESPONSIVE-TO-ANTAGONIST1, a Menkes/Wilson disease-related copper transporter, is required for ethylene signaling in *Arabidopsis*. Cell **97(3)**:383-393

**Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD, Gitlin JD** (1997) Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. J Biol Chem **272(34)**:21461-21466

**Kanamaru K, Kashiwagi S, Mizuno T** (1994) A copper-transporting P-type ATPase found in the thylakoid membrane of the cyanobacterium *Synechococcus* species PCC7942. Mol Microbiol **13(2)**:369-377

**Odermatt A, Suter H, Krapf R, Solioz M** (1993) Primary structure of two P-type ATPases involved in copper homeostasis in *Enterococcus hirae*. J Biol Chem **268(17)**:12775-12779

**Palmgren MG, Harper JF** (1999) Pumping with plant P-type ATPases. J Exp Bot **50**:883-893

**Petris MJ, Mercer JF, Culvenor JG, Lockhart P, Gleeson PA, Camakaris J** (1996) Ligand-regulated transport of the Menkes copper P-type ATPase efflux pump from the Golgi apparatus to the plasma membrane: a novel mechanism of regulated trafficking. EMBO J **15(22)**:6084-6095

**Rensing C, Sun Y, Mitra B, Rosen BP** (1998) Pb(II)-translocating P-type ATPases. J Biol Chem **273(49)**:32614-32617

**Rutherford JC, Cavet JS, Robinson NJ** (1999) Cobalt-dependent transcriptional switching by a dual-effector MerR-like protein regulates a cobalt-exporting variant CPx-type ATPase. J Biol Chem **274(36)**:25827-25832

**Schaefer M, Hopkins RG, Failla ML, Gitlin JD (1999)** Hepatocyte-specific localization and copper-dependent trafficking of the Wilson's disease protein in the liver. Am J Physiol **276**:G639-646

**Solioz M, Odermatt A** (1995) Copper and silver transport by CopB-ATPase in membrane vesicles of *Enterococcus hirae*. J Biol Chem **270(16)**:9217-9221

**Solioz M, Vulpe C** (1996) CPx-type ATPases: a class of P-type ATPases that pump heavy metals. Trends Biochem Sci **21(7)**:237-241

**Tabata K, Kashiwagi S, Mori H, Ueguchi C, Mizuno T** (1997) Cloning of a cDNA encoding a putative metal-transporting P-type ATPase from *Arabidopsis thaliana*. Biochim Biophys Acta **1326(1)**:1-6

**Tottey S, Rich PR, Rondet SA, Robinson NJ** (2001) Two Menkes-type ATPases supply copper for photosynthesis in *Synechocystis* PCC 6803. J Biol Chem **276(23)**:19999-20004

**Tsai KJ, Yoon KP, Lynn AR** (1992) ATP-dependent cadmium transport by the *cadA* cadmium resistance determinant in everted membrane vesicles of *Bacillus subtilis*. J Bacteriol **174(1)**:116-121

**Williams LE, Pittman JK, Hall JL** (2000) Emerging mechanisms for heavy metal transport in plants. Biochim Biophys Acta **1465(1-2)**:104-126

**Woeste KE, Kieber JJ (2000)** A strong loss-of-function mutation in *RAN1* results in constitutive activation of the ethylene response pathway as well as a rosette-lethal phenotype. Plant Cell **12(3)**:443-455

**Yamaguchi Y, Heiny ME, Suzuki M, Gitlin JD** (1996) Biochemical characterization and intracellular localization of the Menkes disease protein. Proc Natl Acad Sci U S A **93(24)**:14030-14035

**Yuan DS, Dancis A, Klausner RD** (1997) Restriction of copper export in *Saccharomyces cerevisiae* to a late Golgi or post-Golgi compartment in the secretory pathway. J Biol Chem **272(41)**:25787-25793

# 5 Appendix

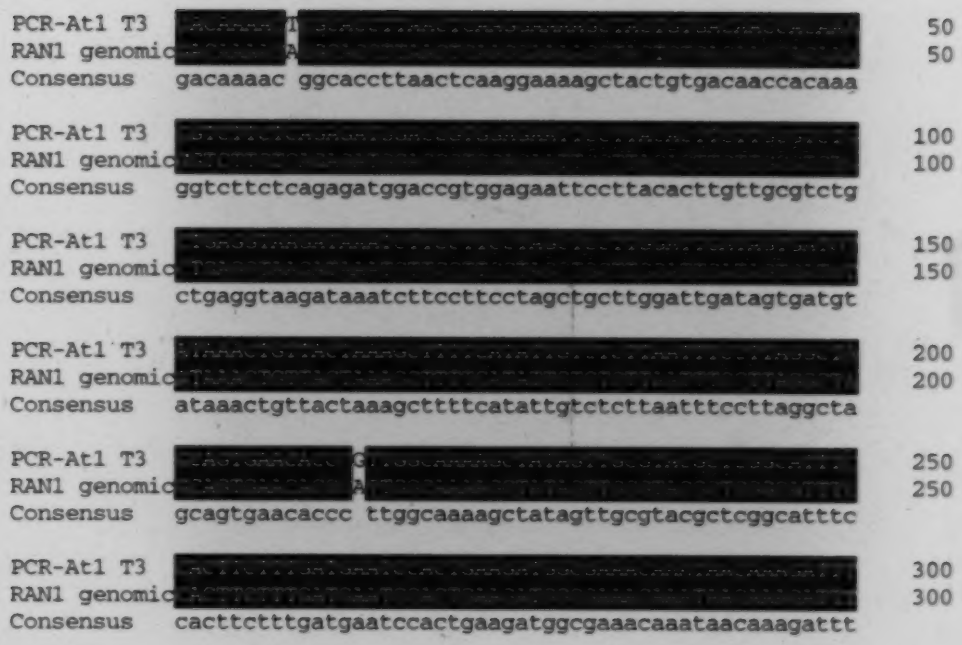


**Figure 5.1 The partial sequence of PCR-At1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* genomic sequence.** The PCR-At1 product, which was produced using the PCR-S/PCR-A primer pair and *Arabidopsis* genomic DNA, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic sequence.

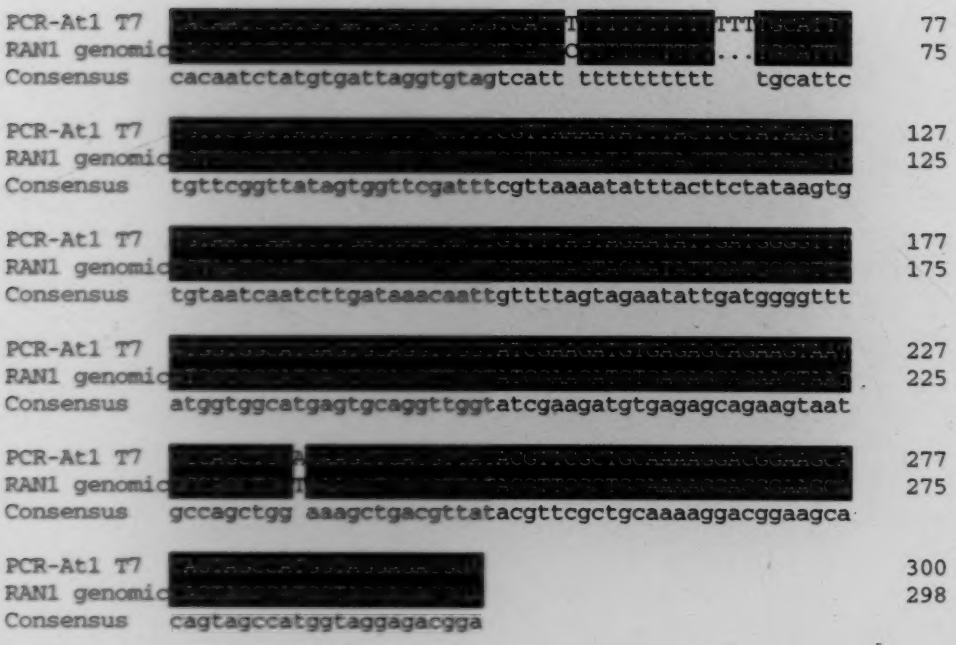


**Figure 5.2 The partial sequence of PCR-At1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* genomic sequence.** The PCR-At1 product, which was produced using the PCR-S/PCR-A primer pair and *Arabidopsis* genomic DNA, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic sequence.

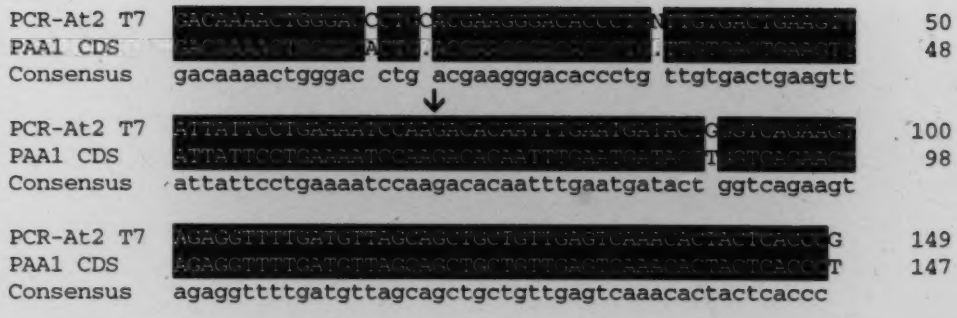


**Figure 5.3 The partial coding sequence (CDS) of PCR-At2 that was produced with the T7 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.** The PCR-At2 product, which was produced using the PCR-S/PCR-A primer pair and *Arabidopsis* genomic DNA, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction contained 151 nucleotides that correspond to a putative intron. Since the *PAA1* genomic sequence was not available, this putative intron sequence was removed and the remaining nucleotides were aligned with the appropriate region of the *PAA1* mRNA/CDS. The arrow (↓) indicates where the putative intron was located.

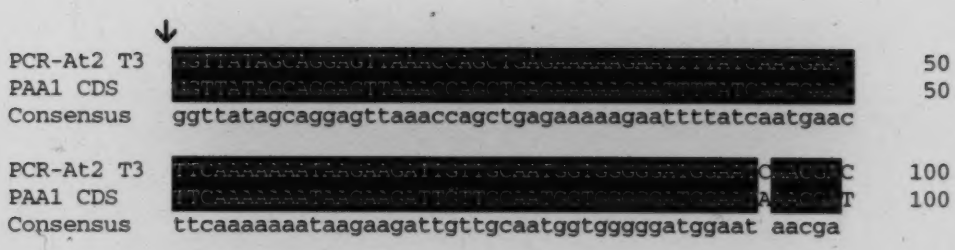


**Figure 5.4 The partial coding sequence (CDS) of PCR-At2 that was produced with the T3 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.** The PCR-At2 product, which was produced using the PCR-S/PCR-A primer pair and *Arabidopsis* genomic DNA, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction contained 200 nucleotides that correspond to a putative intron. Since the *PAA1* genomic sequence was not available, this putative intron sequence was removed and the remaining nucleotides were aligned with the appropriate region of the *PAA1* mRNA/CDS. The arrow (↓) indicates where the putative intron was located.

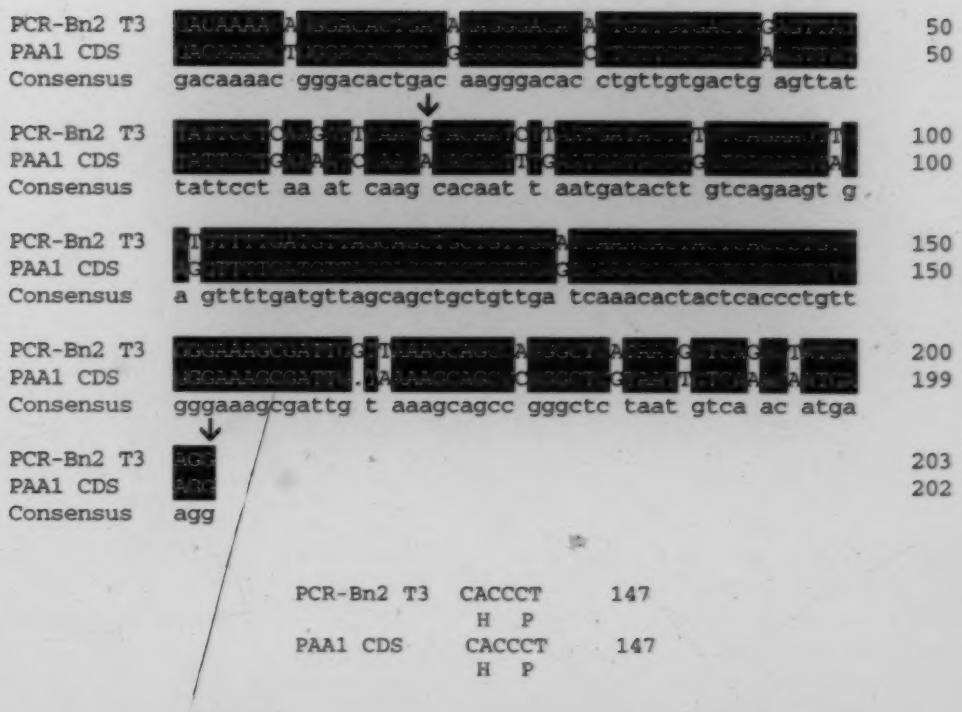


**Figure 5.5 The partial coding sequence (CDS) of PCR-Bn2 that was produced with the T3 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.** The PCR-Bn2 product, which was produced using the PCR-S/PCR-A primer pair and *B. napus* genomic DNA, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction contained 97 nucleotides that correspond to putative introns. Since the *PAA1* genomic sequence was not available, these putative intron sequences were removed and the remaining nucleotides were aligned with the appropriate region of the *PAA1* mRNA/CDS. The arrows (↓) indicate where the putative introns were located. The nucleotides that encode the HP locus are presented below the alignment.

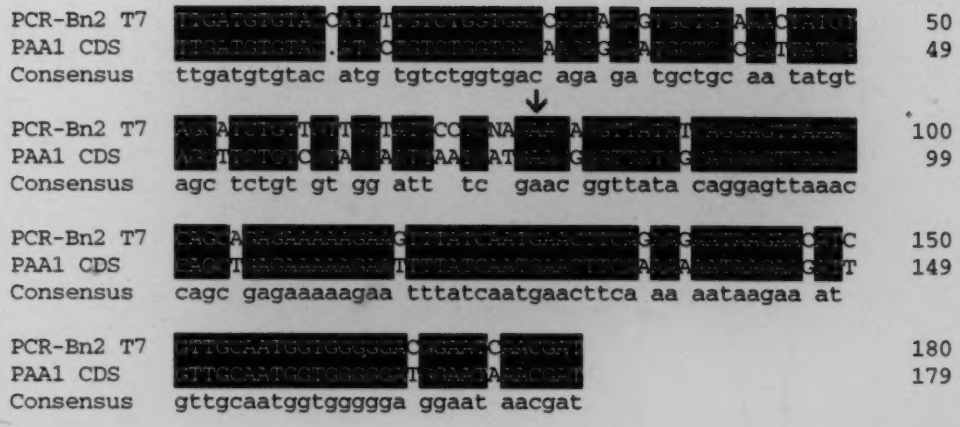


**Figure 5.6 The partial coding sequence (CDS) of PCR-Bn2 that was produced with the T7 primer aligned with the appropriate region of the *PAA1* mRNA/CDS.** The PCR-Bn2 product, which was produced using the PCR-S/PCR-A primer pair and *Brassica napus* genomic DNA, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction contained 120 nucleotides that correspond to a putative intron. Since the *PAA1* genomic sequence was not available, these putative intron sequences were removed and the remaining nucleotides were aligned with the appropriate region of the *PAA1* mRNA/CDS. The arrow (↓) indicates where the putative intron was located.

**Figure 5.7 The partial sequence of the ~2000 bp cDNA that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The ~2000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* coding sequence.

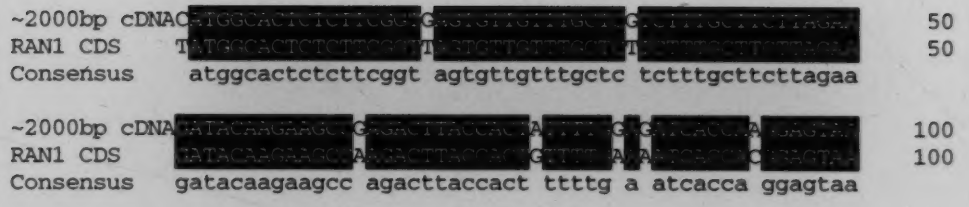


**Figure 5.8 The partial sequence of the ~2000 bp cDNA that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The ~2000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T3 primer.. The first 300 nucleotides from this sequencing reaction contained 200 nucleotides that were from the 3'-untranslated region. In order to compare the sequencing results to the *RAN1* coding sequence, the 3'-untranslated region was removed.

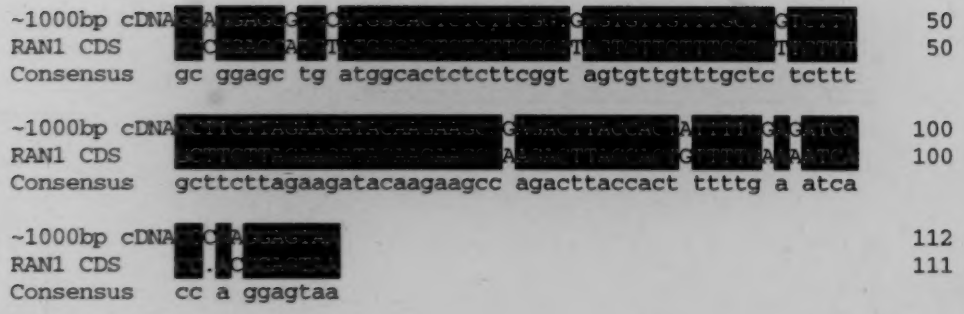


**Figure 5.9 The partial sequence of the ~1000 bp cDNA that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The ~1000 bp cDNA that was isolated from the amplified *B. napus* cDNA library was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction contained 188 nucleotides that were from the 3'-untranslated region. In order to compare the sequencing results to the *RAN1* coding sequence, the 3'-untranslated region was removed.

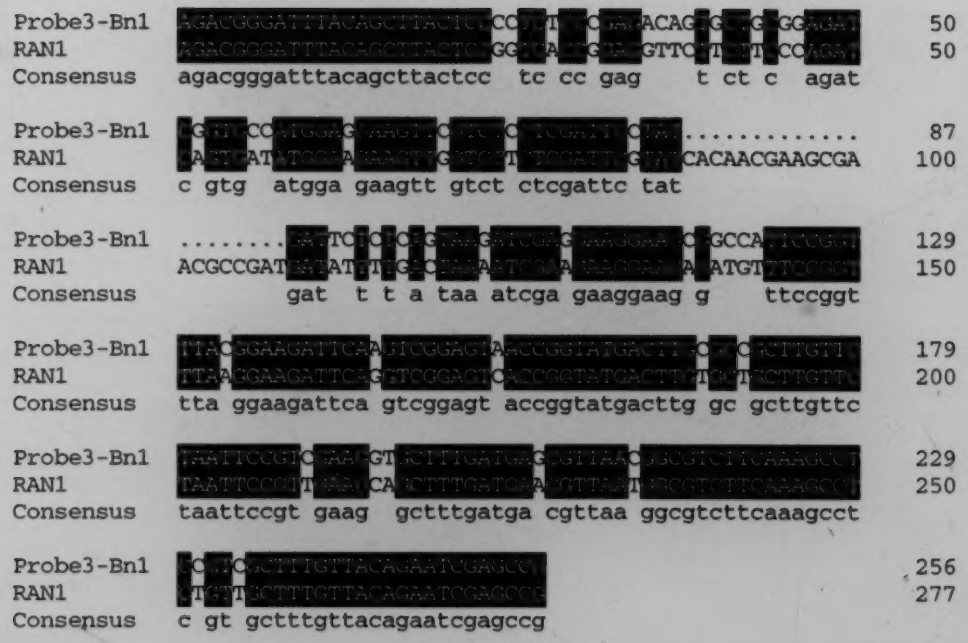


**Figure 5.10 The Probe3-Bn1 sequence aligned with the appropriate region of the *RAN1* genomic or coding sequence.** The Probe3-Bn1 product, which was produced using the Probe3-S/Probe3-A primer pair and *B. napus* genomic DNA, was completely sequenced using the T7 primer. The nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic or coding sequence. There are no introns in the Probe3-Bn1 product and aligning the Probe3-Bn1 sequence with the appropriate region of the *RAN1* genomic and coding sequences produces equivalent results.

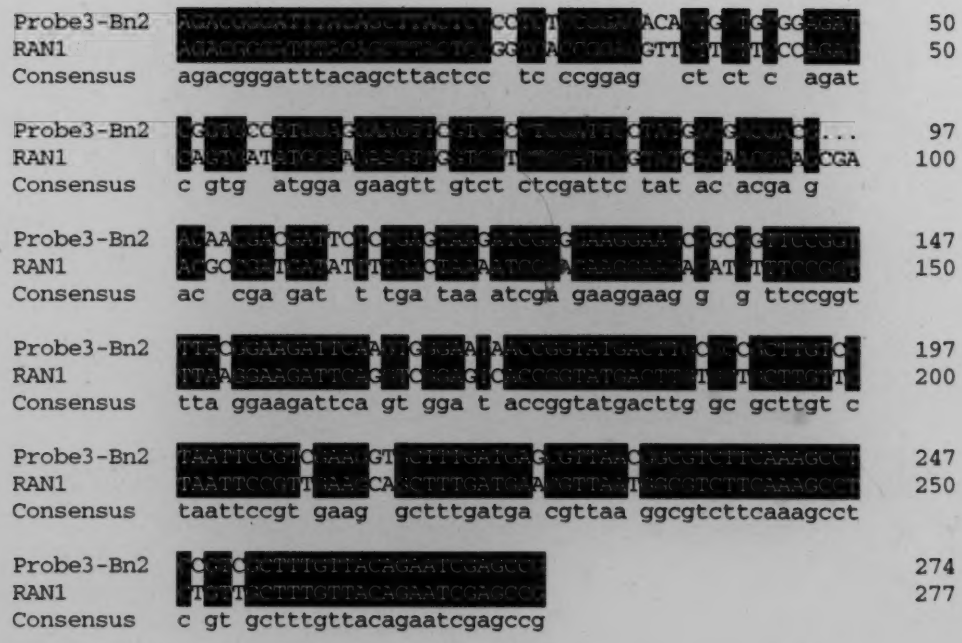


**Figure 5.11 The Probe3-Bn2 sequence aligned with the appropriate region of the *RAN1* genomic or coding sequence.** The Probe3-Bn2 product, which was produced using the Probe3-S/Probe3-A primer pair and *B. napus* genomic DNA, was completely sequenced using the T7 primer. The nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* genomic or coding sequence. There are no introns in the Probe3-Bn2 product and aligning the Probe3-Bn2 sequence with the appropriate region of the *RAN1* genomic and coding sequences produces equivalent results.

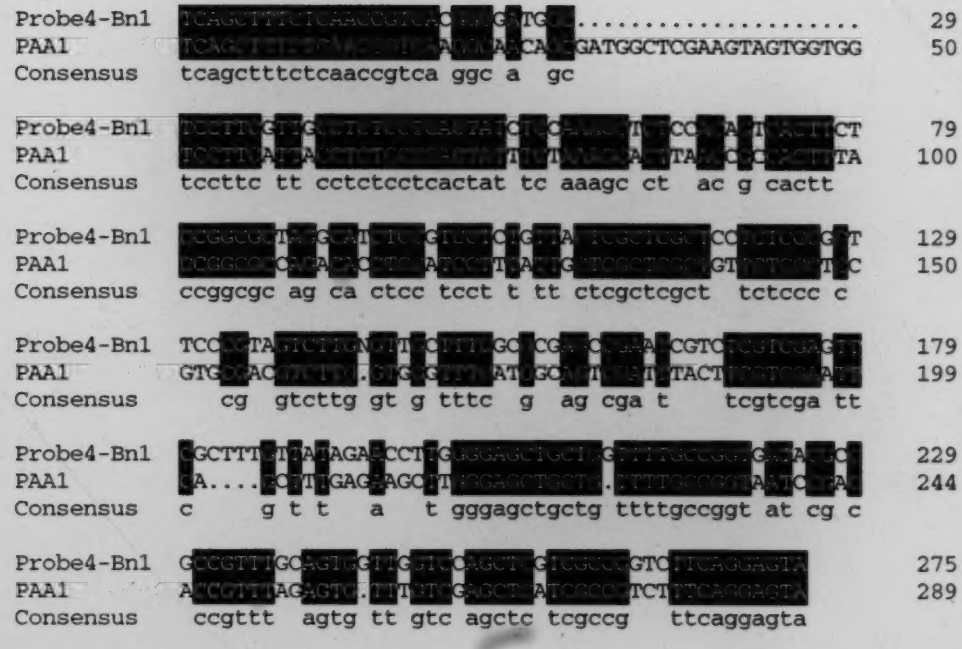


**Figure 5.12 The Probe4-Bn1 sequence aligned with the appropriate region of the *PAA1* genomic or mRNA/CDS.** The Probe4-Bn1 product, which was produced using the Probe4-S/Probe4-A primer pair and *B. napus* genomic DNA, was completely sequenced using the T3 primer. The nucleotides from this sequencing reaction were aligned with the appropriate region of the *PAA1* genomic or mRNA/CDS. There are no introns in the Probe4-Bn1 product and aligning the Probe4-Bn1 sequence with the appropriate region of the *PAA1* genomic sequence and mRNA/CDS produces equivalent results.

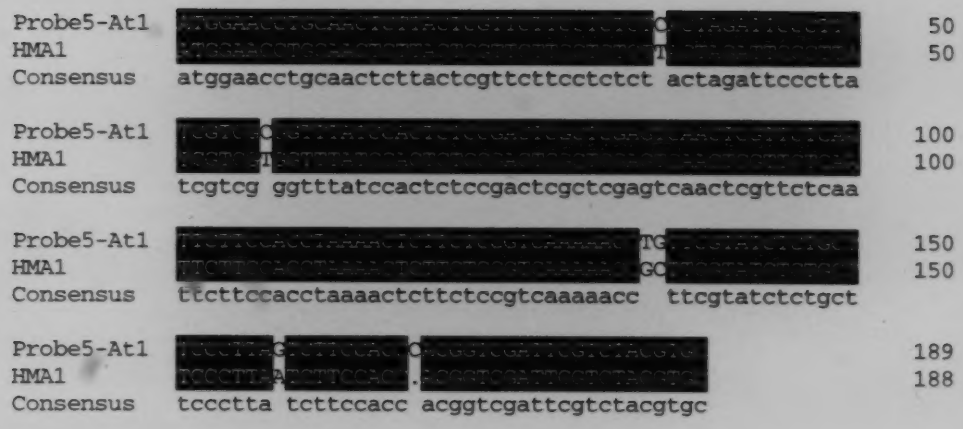


**Figure 5.13 The Probe5-At1 sequence aligned with the appropriate region of the *HMA1* genomic or coding sequence.** The Probe5-At1 product, which was produced using the Probe5-S/Probe5-A primer pair and *Arabidopsis* genomic DNA, was completely sequenced using the T7 primer. The nucleotides from this sequencing reaction were aligned with the appropriate region of the *HMA1* genomic or coding sequence. There are no introns in the Probe5-At1 product and aligning the Probe5-At1 sequence with the appropriate region of the *HMA1* genomic and coding sequences produces equivalent results.

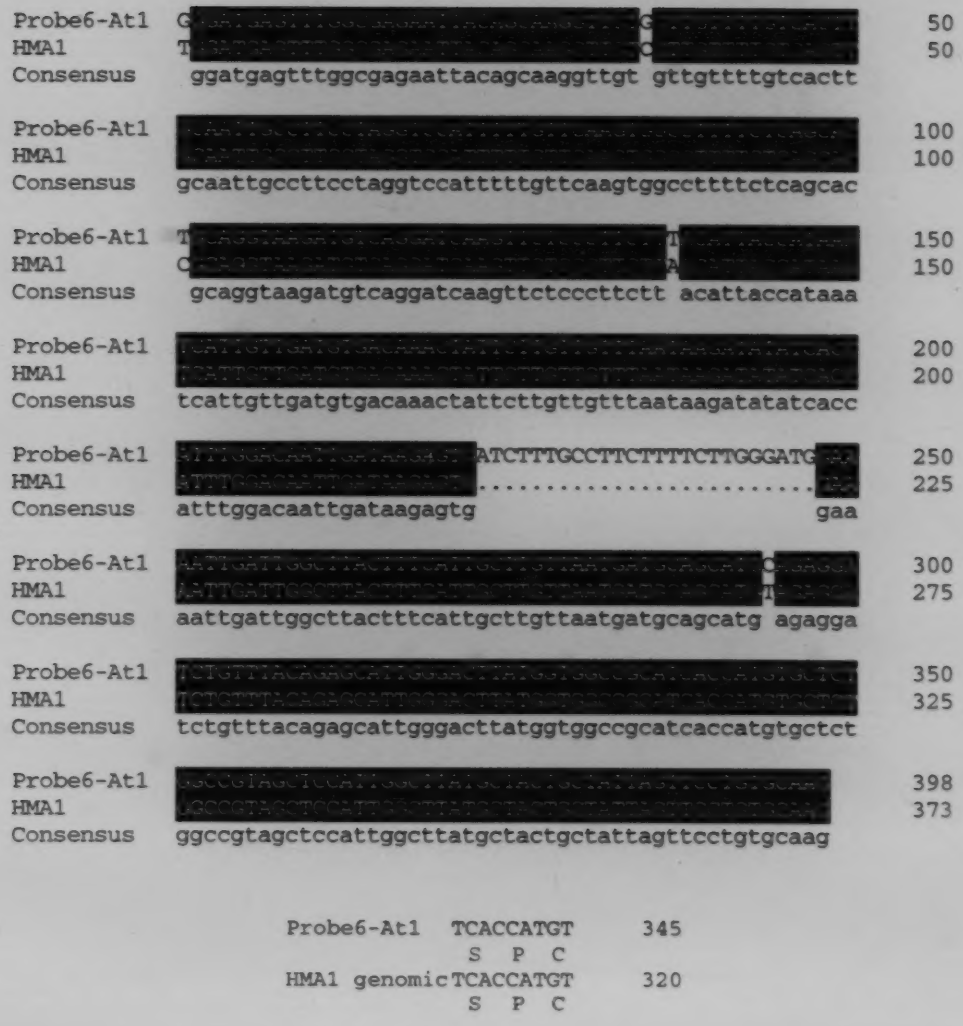


**Figure 5.14 The Probe6-At1 sequence aligned with the appropriate region of the *HMA1* genomic sequence.** The Probe6-At1 product, which was produced using the Probe6-S/Probe6-A primer pair and *Arabidopsis* genomic DNA, was completely sequenced using the T3 and T7 primers. The Probe6-At1 sequence was aligned with the appropriate region of the *HMA1* genomic sequence. A translation is presented for the portion of the alignment that contained the region of the *HMA1* genomic sequence that encodes the SPC tripeptide.

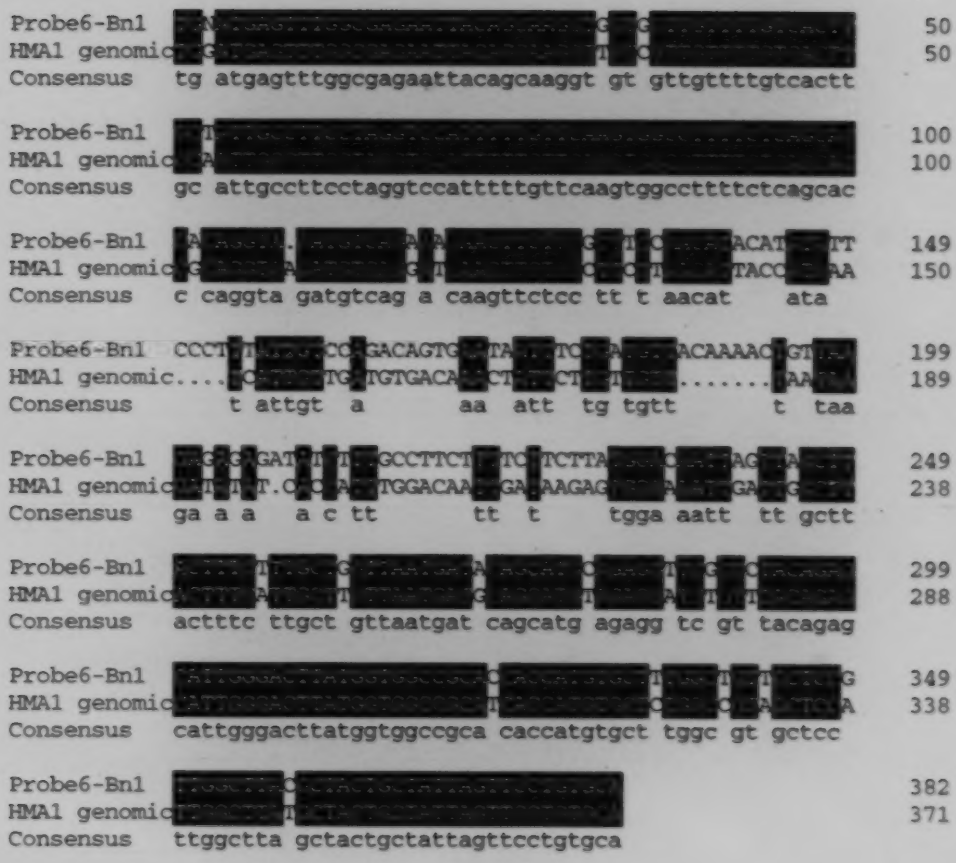


**Figure 5.15 The Probe6-Bn1 sequence aligned with the appropriate region of the *HMA1* genomic sequence.** The Probe6-Bn1 product, which was produced using the Probe6-S/Probe6-A primer pair and *B. napus* genomic DNA, was completely sequenced using the T3 and T7 primers.

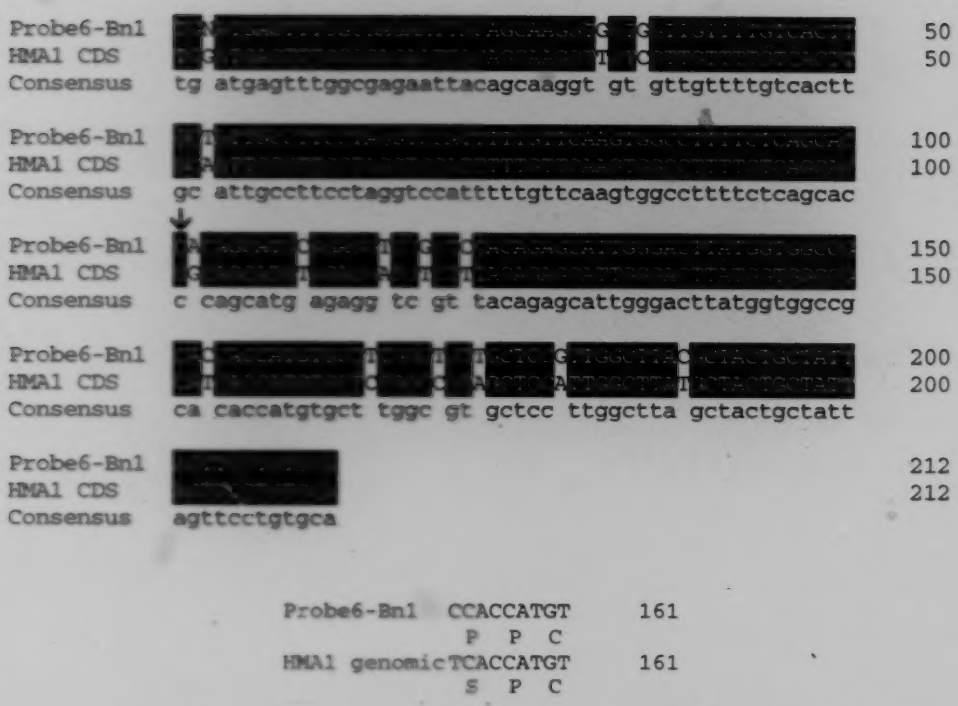


**Figure 5.16 The Probe6-Bn1 coding sequence (CDS) aligned with the appropriate region of the *HMA1* coding sequence.** The Probe6-Bn1 product, which was produced using the Probe6-S/Probe6-A primer pair and *B. napus* genomic DNA, was completely sequenced using the T3 and T7 primers. The Probe6-Bn1 product contained 170 nucleotides that correspond to a putative intron. In order to compare the sequencing results to the *HMA1* coding sequence, this putative intron sequence was removed. The arrow (↓) indicates where the putative intron would be located. A translation is presented for the portion of the alignment that contained the region of the *HMA1* coding sequence that encodes the SPC tripeptide.

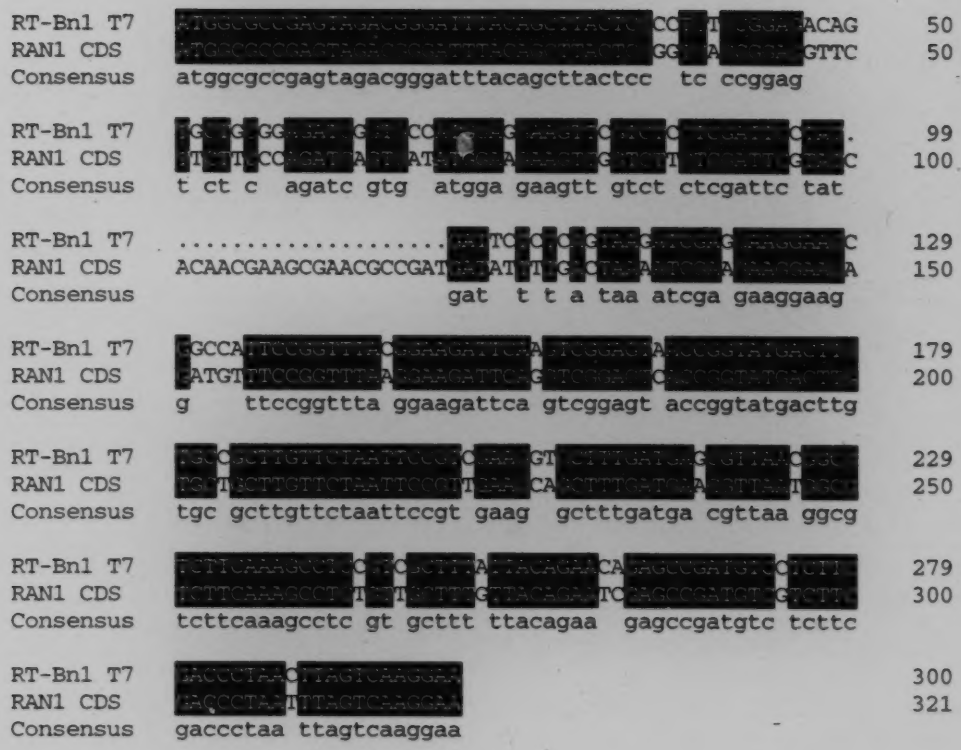


**Figure 5.17 The partial coding sequence of RT-Bn1 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The RT-Bn1 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the RT3-A primer, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* coding sequence (CDS).

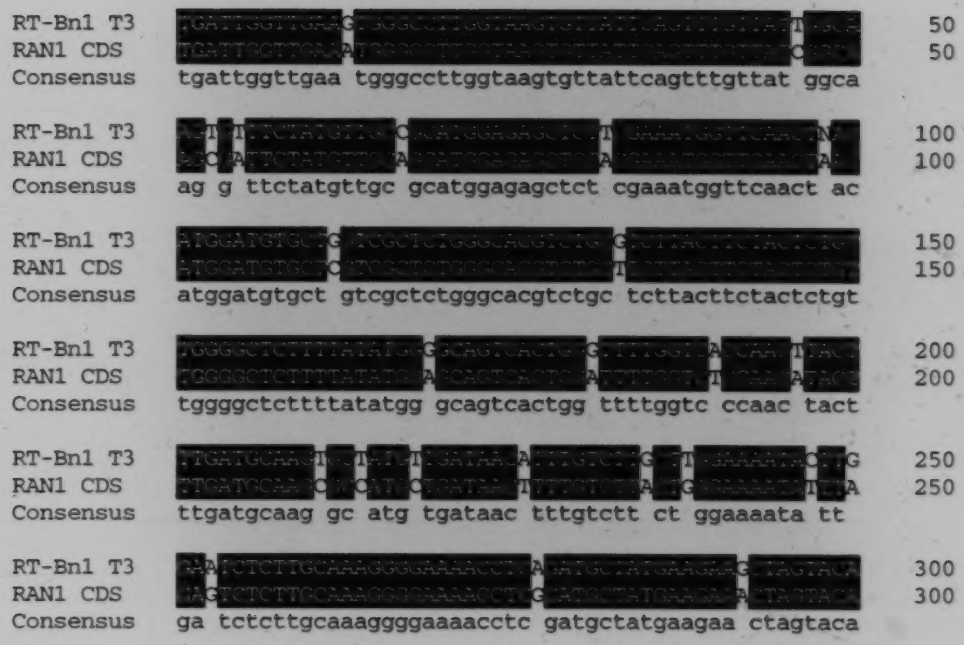


**Figure 5.18 The partial coding sequence of RT-Bn1 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The RT-Bn1 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the RT3-A primer, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* coding sequence (CDS).

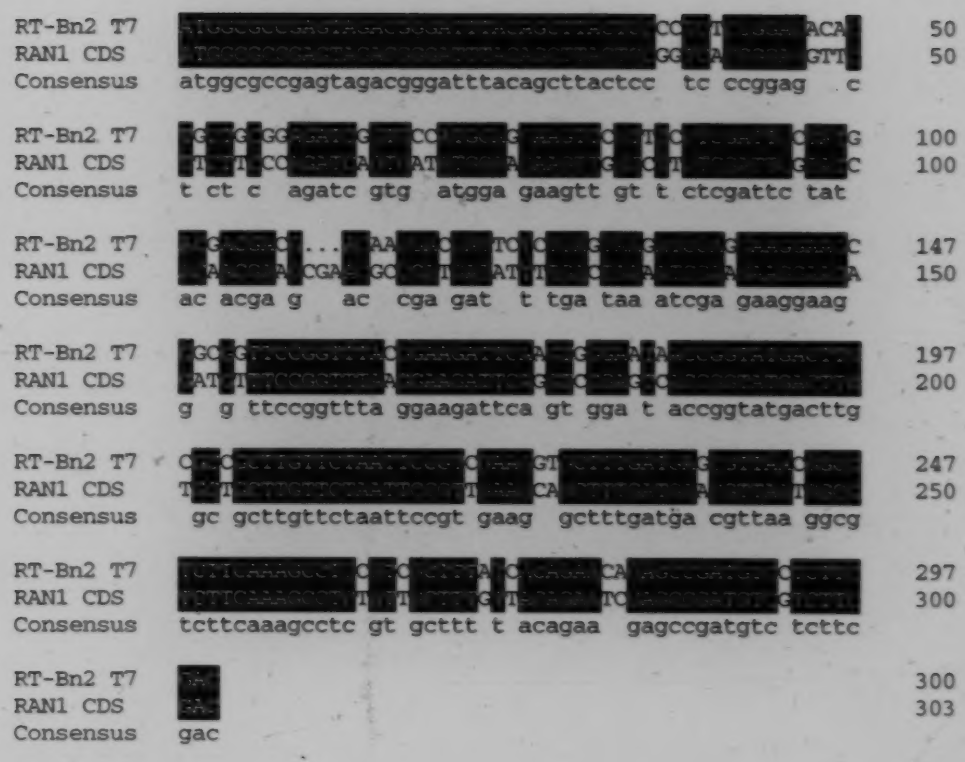


**Figure 5.19 The partial coding sequence of RT-Bn2 that was produced with the T7 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The RT-Bn2 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the Probe1-A primer, was partially sequenced using the T7 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* coding sequence (CDS).

**Figure 5.20 The partial coding sequence of RT-Bn2 that was produced with the T3 primer aligned with the appropriate region of the *RAN1* coding sequence (CDS).** The RT-Bn2 product, which was produced using the RT2-S / RT1/2-A primer pair and first strand cDNA produced from *B. napus* total RNA using the Probe1-A primer, was partially sequenced using the T3 primer. The first 300 nucleotides from this sequencing reaction were aligned with the appropriate region of the *RAN1* coding sequence (CDS).